

Speeded up robust features

From Wikipedia, the free encyclopedia

In computer vision, **Speeded Up Robust Features (SURF)** is a local feature detector and descriptor 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.

To detect interest points, SURF uses an integer approximation of the determinant of Hessian blob detector, which can be computed with 3 integer operations using a precomputed integral image. Its feature descriptor is based on 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, to make 3D scenes, to track objects and to 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 United States.^[1]

Contents

- 1 Overview
- 2 Algorithm and features
 - 2.1 Interest point detection
 - 2.1.1 Scale-space representation and location of points of interest
 - 2.2 Local neighborhood descriptor
 - 2.2.1 Orientation assignment
 - 2.2.2 Descriptor based on the sum of Haar wavelet responses
 - 2.3 Matching
- 3 See also
- 4 References
- 5 External links

Overview

SURF is a detector and a descriptor for points of interest in images where the image is transformed into coordinates, using the multi-resolution pyramid technique. Is to make a copy of the original image with Pyramidal Gaussian or Laplacian Pyramid shape and obtain image with the same size but with reduced bandwidth. Thus a special blurring effect on the original image, called Scale-Space is achieved. This technique ensures that the points of interest are scale invariant.

Algorithm and features

The SURF algorithm is based on the same principles and steps as SIFT but details in each step are different. The algorithm has three main parts: interest point detection, local neighborhood description and matching.

Interest point detection

The SIFT approach uses cascaded filters to detect scale-invariant characteristic points, where the difference of Gaussians (DoG) is calculated on rescaled images progressively. In SURF, square-shaped filters are used as an approximation of Gaussian smoothing₇. Filtering the image with a square is much faster if the integral image₇ is used, which is defined as:

$$S(x, y) = \sum_{i=0}^x \sum_{j=0}^y I(i, j)$$

The sum of the original image within a rectangle can be evaluated quickly using the integral image, requiring four evaluations at the corners of the rectangle.

SURF uses a blob detector based on the Hessian matrix₇ to find points of interest. The determinant₇ of the Hessian matrix is used as a measure of local change around the point and points are chosen where this determinant is maximal. In contrast to the Hessian-Laplacian detector by Mikolajczyk and Schmid, SURF also uses the determinant of the Hessian for selecting the scale, as it is done by Lindeberg. Given a point $p=(x, y)$ in an image I , the Hessian matrix $H(p, \sigma)$ at point and scale σ , is defined as follows:

$$H(p, \sigma) = \begin{pmatrix} L_{xx}(p, \sigma) & L_{xy}(p, \sigma) \\ L_{xy}(p, \sigma) & L_{yy}(p, \sigma) \end{pmatrix}$$

where $L_{xx}(p, \sigma)$ etc. are the second-order derivatives of the grayscale image.

The box filter of size 9×9 is an approximation of a Gaussian with $\sigma=1.2$ and represents the lowest level (highest spatial resolution) for blob-response maps.

Scale-space representation and location of points of interest

The interest points can be found in different scales, partly because the search for correspondences often requires comparison images where they are seen at different scales. The scale space is usually realized as an image pyramid. Images are repeatedly smoothed with a Gaussian filter, then they are subsampled to get the next higher level of the pyramid. Therefore, several floors or stairs with various measures of the masks are calculated:

$$\sigma_{approx} = Currentfiltersize * \left(\frac{(BaseFilterscale)}{BaseFilterSize} \right)$$

The scale space is divided into a number of octaves, where an octave refers to a series of response maps of covering a doubling of scale. In SURF, the lowest level of the scale space is obtained from the output of the 9×9 filters.

Scale spaces are implemented by applying box filters of different sizes. Therefore, the scale space is analyzed by up-scaling the filter size rather than iteratively reducing the image size. The output of the above 9×9 filter is considered as the initial scale layer, to which we will refer as scale $s=1.2$ (corresponding to Gaussian derivatives with $\sigma=1.2$). The following layers are obtained by filtering the image with gradually bigger masks, taking into account the discrete nature of integral images and the specific structure of filters. Specifically, this results in filters of size 9×9, 15×15, 21×21, 27×27, etc. In order to localize interest points in the image and over scales, non-maximum suppression in a 3×3×3 neighborhood is applied. The maxima of the determinant of the Hessian matrix are then interpolated in scale and image space with the method proposed by Brown et al. Scale space interpolation is especially important in this case, as the difference in scale between the first layers of every octave is relatively large.

Local neighborhood descriptor

The goal of a descriptor is to provide a unique and robust description of an image feature[?], e.g. by describing the intensity distribution of the pixels within the neighborhood of the point of interest. Most descriptors are computed thus in a local manner; hence, a description is obtained for every point of interest identified previously.

The dimensionality of the descriptor has direct impact on both its computational complexity and point-matching robustness/accuracy. A short descriptor may be more robust against appearance variations, but may not offer sufficient discrimination and thus give too many false positives.

The first step consists of fixing a reproducible orientation based on information from a circular region around the interest point. Then we construct a square region aligned to the selected orientation and extract the SURF descriptor from it.

Furthermore, there is also an upright version of SURF (called U-SURF) that is not invariant to image rotation and therefore faster to compute and better suited for application where the camera remains more or less horizontal.

Orientation assignment

In order to achieve rotational invariance, the orientation of the point of interest needs to be found. The Haar wavelet responses in both x- and y-directions within a circular neighbourhood of radius $6s$ around the point of interest are computed, where s is the scale at which the point of interest was detected. The obtained responses are weighed by a Gaussian function centered at the point of interest, then plotted as points in a two-dimensional space, with the horizontal response in the abscissa and the vertical response in the ordinate. The dominant orientation is estimated by calculating the sum of all responses within a sliding orientation window of size $\pi/3$. The horizontal and vertical responses within the window are summed. The two summed responses then yield a local orientation vector. The longest such vector overall defines the orientation of the point of interest. The size of the sliding window is a parameter which has to be chosen carefully to achieve a desired balance between robustness and angular resolution.

Descriptor based on the sum of Haar wavelet responses

To describe the region around the point, a square region is extracted centred around the interest point and oriented along the orientation as selected in the previous section. The size of this window is $20s$.

The interest region is split up into smaller 4×4 square sub-regions, and for each one, the Haar wavelet responses are extracted at 5×5 regularly spaced sample points. The responses are weighted with a Gaussian (to offer more robustness for deformations, noise and translation).

Matching

By comparing the descriptors obtained from different images, matching pairs can be found.

See also

- Scale-invariant feature transform[?] (SIFT)
- Gradient Location and Orientation Histogram[?]
- LESH - Local Energy based Shape Histogram[?]
- Blob detection[?]
- Feature detection (computer vision)[?]

References

1. US 2009238460 (<http://worldwide.espacenet.com/textdoc?DB=EPODOC&IDX=US2009238460>), Ryuji Funayama, Hiromichi Yanagihara, Luc Van Gool, Tinne Tuytelaars, Herbert Bay, "ROBUST INTEREST POINT DETECTOR AND DESCRIPTOR", published 2009-09-24
- Herbert Bay, Tinne Tuytelaars, and Luc Van Gool, "Speeded Up Robust Features (<http://www.vision.ee.ethz.ch/~surf/eccv06.pdf>)", ETH Zurich, Katholieke Universiteit Leuven
 - Andrea Maricela Plaza Cordero, Jorge Luis Zambrano Martínez, " Estudio y Selección de las Técnicas SIFT, SURF y ASIFT de Reconocimiento de Imágenes para el Diseño de un Prototipo en Dispositivos Móviles (http://41jaiio.sadio.org.ar/sites/default/files/6_EST_2012.pdf)" , 15° Concurso de Trabajos Estudiantiles, EST 2012
 - A. M. Romero and M. Cazorla, "Comparativa de detectores de características visuales y su aplicación al SLAM (http://rua.ua.es/dspace/bitstream/10045/23389/1/2009_Romero_Cazorla_Workshop_Agentes_Fisicos_Caceres.pdf) ", X Workshop de agentes físicos, Setiembre 2009, Cáceres
 - P. M. Panchal, S. R. Panchal, S. K. Shah, "A Comparison of SIFT and SURF (http://www.ijircce.com/upload/2013/april/21_V1204057_A%20Comparison_H.pdf) ", International Journal of Innovative Research in Computer and Communication Engineering Vol. 1, Issue 2, April 2013
 - Herbert Bay, Andreas Ess, Tinne Tuytelaars, Luc Van Gool "SURF: Speeded Up Robust Features" (<http://www.vision.ee.ethz.ch/~surf/papers.html>), Computer Vision and Image Understanding (CVIU), Vol. 110, No. 3, pp. 346–359, 2008
 - Christopher Evans "Notes on the OpenSURF Library", MSc Computer Science, University of Bristol (<http://www.chrisevansdev.com>)

External links

- Website of SURF: Speeded Up Robust Features (<http://www.vision.ee.ethz.ch/~surf>)
- First publication of Speeded Up Robust Features (2006) (<http://www.vision.ee.ethz.ch/~surf/eccv06.pdf>)
- Revised publication of SURF (2008) (http://glorfindel.mavrinac.com/~aaron/school/pdf/bay06_surf.pdf)

Retrieved from "http://en.wikipedia.org/w/index.php?title=Speeded_up_robust_features&oldid=663739650"

Categories : Feature detection (computer vision)

-
- This page was last modified on 24 May 2015, at 01:16.
 - Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. By using this site, you agree to the Terms of Use and Privacy Policy. Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a non-profit organization.