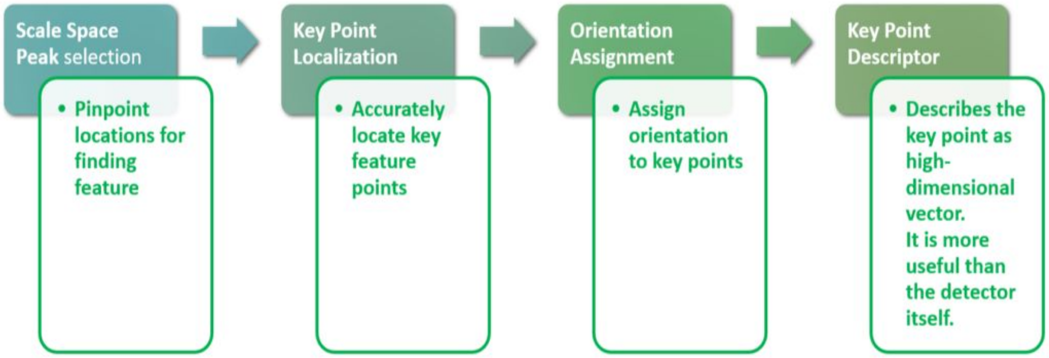
Feature Extraction Methods

Group Dynamic

SIFT (Scale-Invariant Feature Transform)

- SIFT is used in the detection of interest points on an input image and it allows identification of localized features in images.It permits distinguishing proof of localized highlights in pictures which is fundamental in applications like; Object Recognition in Images,Path detection and obstacle avoidance algorithms, and Gesture recognition, Mosaic generation, etc
- SIFT can perform feature detection independent of these properties (viewpoints, depths, and scale) of the image.
- SIFT detector follows four sequence of steps which are;

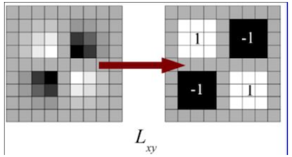


SURF (Speeded Up Robust Features)

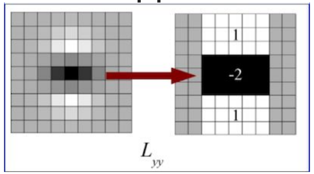
- Introduced by Herbert Bay, Tinne Tuytelaars, and Luc Van Gool in 2006 to offer a more rapid model than the SIFT model.
- Model consists of two main steps:

1. Orientation assignment: First, interest point is detected by Hessian matrix approximation using box filters.

Gaussian partial derivative in xy:



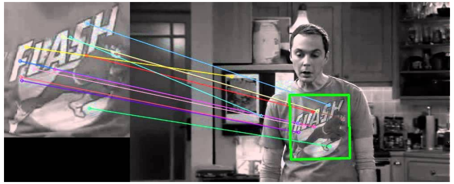
in y:



Then, the model fixes a reproducible orientation based on information from a circular region around the interest point in order to be invariant to rotation.

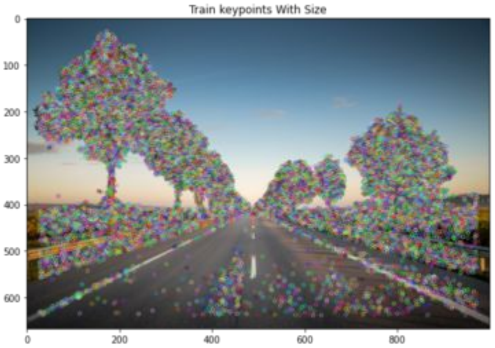
2. Feature description: First, square region around the interest point along the selected orientation is created.

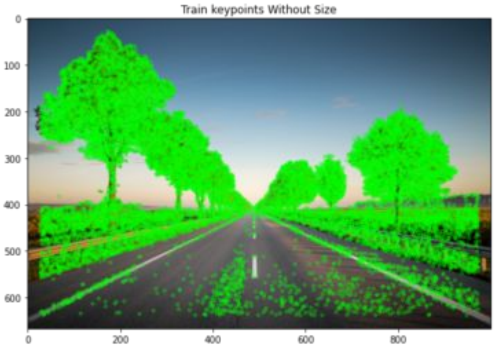
Then, the descriptor is extracted using this square region with the help of 4*4 sub-regions and 5*5 sample points within them.



BRIEF (Binary Robust Independent Elementary Features)

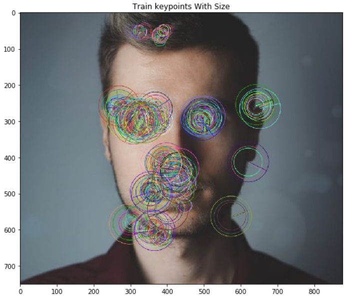
- BRIEF is merely a feature descriptor that can be used in conjunction with any other feature detector. By translating descriptors in floating point integers to binary strings, this approach minimizes memory utilization.
- One thing to keep in mind is that BRIEF is a feature descriptor, not a method for finding features. As a result, you will have to rely on other feature detectors like SIFT, SURF, and so on.
- Shortly, BRIEF is a quicker way for calculating and matching feature descriptors. Unless there is a lot of in-plane rotation, it also has a high recognition rate.

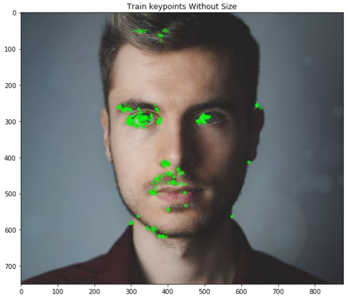




ORB (Oriented FAST and rotated BRIEF)

- ORB is based on the FAST keypoint detector and a modified version of the visual descriptor BRIEF.
- A learning method for decorrelating BRIEF features under rotational invariance, leading to better performance in nearest-neighbor applications.
- Analysis of variance and correlation of oriented BRIEF features
- An efficient and viable alternative to SIFT and SURF. ORB was conceived mainly because SIFT and SURF are patented algorithms. However, ORB is free to use.
- ORB performs as well as SIFT on the task of feature detection (and is better than SURF) while being almost two orders of magnitude faster.





Comparison

- BRIEF is faster than SIFT in terms of rapidity and accuracy rate.
- BRIEF reduces memory usage as it converts descriptor in float to binary strings.
- SURF is faster than SIFT.
- ORB is the fastest algorithm while SIFT performs the best in the most scenarios.
- SIFT and SURF are patented so not free for commercial use, while ORB is free. SIFT and SURF detect more features then ORB, but ORB is faster.

Descriptor	Type	Detector	Sampling Pairs	Rotation Invariant	Scale Invariant
SIFT	Float-type	SIFT	NA	YES	YES
SURF	Float-type	SURF	NA	YES	YES
BRIEF	Binary	SURF	Random	NO	NO
ORB	Binary	ORB	Learned	YES	NO

Resources

https://docs.opencv.org/3.4/df/dd2/tutorial_py_surf_intro.html

https://www.geeksforgeeks.org/sift-interest-point-detector-using-python-opencv/

https://medium.com/data-breach/introduction-to-orb-oriented-fast-and-rotated-brief-4220e8ec40cf

Bay, H., Tuytelaars, T., & Gool, L. V. (2006, May). Surf: Speeded up robust features. *European conference on computer vision* (pp. 404-417). Springer, Berlin, Heidelberg.

https://medium.com/data-breach/introduction-to-surf-speeded-up-robust-features-c7396d6e7c4e

Hassaballah, M. & Alshazly, H. & Ali, A. (2019). Analysis and Evaluation of Keypoint Descriptors for Image Matching. *Studies in Computational Intelligence*.

Feature Selection Methods

Group Dynamic

Feature Importance

Feature importance refers to a set of strategies for assigning scores to input characteristics in a predictive model, indicating the relative importance of each item when producing a prediction.

For issues involving forecasting a numerical value, called regression, and problems involving predicting a class label, called classification, feature importance scores can be generated. Feature importance gives us 3 benefits:

1)Feature importance scores can provide insight into the dataset.

2)Feature importance scores can provide insight into the model.

3)Feature importance can be used to improve a predictive model.

Correlation Matrix with Heatmap

A correlation heatmap shows a 2D correlation matrix between two discrete dimensions, using colored cells to represent data from usually a monochromatic scale. The value of first dimension appear as the rows of the table while of the second dimension as a column.

Correlation states how the features are related to each other or the target variable. Correlation can be positive (increase in one value of feature increases the value of the target variable) or negative (increase in one value of feature decreases the value of the target variable)

Correlation ranges -1 to +1. Values closer to zero means there is no linear between to variables. The close to 1 the correlation is the more positively correlated they are. Close to -1 means that they are strongly correlated to -1.

This is the kind of a filter method.

Exhaustive Feature Selection

In short, in the exhaustive feature selection he performance of a machine learning algorithm is evaluated against all possible combinations of the features in the dataset.

This is one of the most robust feature selection method.Typically a brute-force assessment of each include subset. This implies that it tries each possible combination of the factors and returns the leading performing subset.

Since the exhaustive selection for ideal highlight subset is infeasible in most cases, numerous look techniques have been proposed within the writing. The regular applications of include determination are in classification, clustering, and relapse tasks.

LASSO Regularization (L1)

Regularization is applying a penalty to different features in the model to avoid over-fitting. This penalty is added to coefficients of the features.

In the LASSO regularization, L1, the penalty may cause some of the coefficients to be zero causing the elimination of the feature.

Because of these, LASSO is an embedded feature selection method.

Comparison

Feature selection method	Filter	Wrapper	Embedded
Technique	Statistical measures	Optimization algorithm	Combination of filter and wrapper method
Computational efficiency	Efficient	Inefficient	Inefficient
Computation time	Time efficient	Slow	Slow
Computational cost	Cheaper	Expensive	Expensive
Computational space	Less computational space	More computational space	More computational space
Complexity	Low	High	High
Generality	High	Less	Less
For high dimensional data	Suitable	Not suitable	Not suitable
Advantage		High classification accuracy	Reduce the computation time
Disadvantage	Does not deal with redundant features	Increased runtime	

Resources

https://scikit-learn.org/stable/auto_examples/feature_selection/plot_feature_selection.html#sphx-glr-auto-examples-feature-selection-plot-feature-selection-py

https://www.analyticsvidhya.com/blog/2020/10/feature-selection-techniques-in-machine-learning/

https://stackabuse.com/applying-wrapper-methods-in-python-for-feature-selection/

https://www.analyticsvidhya.com/blog/2021/06/feature-selection-techniques-in-machine-learning-2/

https://thecleverprogrammer.com/2020/06/30/feature-selection-techniques-in-machine-learning-with-python/

https://towardsdatascience.com/feature-selection-techniques-in-machine-learning-with-python-f24e7da3f36e

https://www.researchgate.net/publication/342050177/figure/tbl1/AS:903836717834240@1592502874052/Comparison-of-feature-selection-methods.png