EPFL-TTO

EPFL Innovation Park J
CH-1015 Lausanne
Switzerland
+41 21 693 70 23
http://tto.epfl.ch/

Research Contact

Prof. Pascal Fua
+41 21 693 66 47
pascal.fua@epfl.ch

TTO Contact

Dr Thomas Bourgeau
+41 21 693 3354
thomas.bourgeau@epfl.ch

Method, System, and Device for Learned Invariant Feature Transform for Computer Images

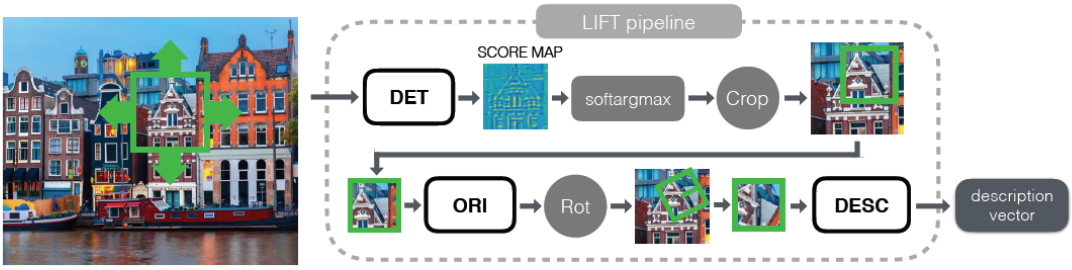

Figure 1: Integrated feature extraction pipeline, consisting of the Detector, Orientation Estimator and Descriptor.

Ref. Nr

6.1672

Keywords

Deep learning, feature detection, feature handling, computer vision, image processing

Intellectual Property

US Patent US 10,552,709 B2

Publications

"LIFT: Learned Invariant Feature Transform" in Computer Vision - Eccv 2016, Pt Vi, 9910, 467-483, 2016

Date

09/11/2020

Description

Detecting and identifying local features in images is an essential part of computer vision software, especially in image matching applications. Recently, deep learning architectures have gained popularity by outperforming traditional methods. However, these new approaches only address a single step in the processing chain of images, which includes feature detection, orientation computation and representation extraction. This technology, referred to as "Learned Invariant Feature Transform" (LIFT), provides a solution to this limitation by performing all these steps together in an optimized way.

LIFT consists of a method for training a feature detector of an image processor. First, a score map is produced by detecting features in the image. The location of the center of mass of the score map is then calculated and used to estimate the orientation of a local patch of the image. Once the orientation of the patch is established, it can be rotated to give a feature description. A Siamese network is used in order to effectively train, in sequence, the descriptor, orientation estimator and feature detector, each relying on the preceding component or components.

Advantages

This invention outperforms the current state-of-the-art on various benchmark data sets, without the need for retraining. It achieves this in part by successfully addressing the complete processing chain and optimizing each stage to work in conjunction with each other.

Training a deep neural network with the widely-used technique of back-propagation requires functions in the network to be differentiable. However, traditional methods of processing the score map, such as non-local maximum suppression (NMS), are not differentiable. LIFT crucially benefits from being able to be trained with back-propagation by reducing the problem into a patch-based setup, and replacing NMS by the so-called 'soft argmax' function, which makes the entire network differentiable.

Applications

- Computer vision, image matching