Kernel Methods for Visual Perception

The dissertation addresses the problem of visual perception, especially in robotics and machine vision. The main task of visual perception is to enable the capability of a computer system to interpret the surrounding environment using data acquired from cameras and other aided sensors. The dissertation proposes the kernel methods for several basic visual perceptual problems, including tracking, localization, mapping, and recognition.

For visual object tracking, this dissertation proposes a kernel cross-correlator (KCC) that breaks traditional limitations. First, by introducing the kernel trick, the KCC extends the linear cross-correlation to non-linear space, which is more robust to signal noises and distortions. Second, the connection to the existing works shows that KCC provides a unified solution for correlation filters. Third, KCC is applicable to any kernel function and is not limited to circulant structure on training data, thus it is able to predict affine transformations with customized properties. Last, by leveraging the fast Fourier transform (FFT), KCC eliminates direct calculation of kernel vectors, thus achieves better performance yet still with a reasonable computational cost. Comprehensive experiments on visual tracking and human activity recognition using wearable devices demonstrate its robustness, flexibility, and efficiency.

Robust velocity and position estimation is crucial for autonomous robot navigation. This dissertation proposes a kernel based algorithm to determine optical flow using a monocular camera, which is named as correlation flow (CF). CF can provide reliable and accurate velocity estimation and is robust to motion blur. In addition, it can also estimate the altitude velocity and yaw rate, which are not available by traditional methods. Autonomous flight tests on a quadcopter show that CF can provide robust trajectory estimation with very low processing power.

In the problem of simultaneous localization and mapping (SLAM), traditional odometry methods resort to iterative algorithms which are usually computationally expensive or require well-designed initialization. To overcome this problem, we propose a non-iterative solution to RGB-D-inertial odometry system, also using the kernel methods. To reduce the odometry and inertial drifts, we present two frameworks for Non-Iterative SLAM (NI-SLAM). One is to seek the aids from ultra wide-band (UWB) technology, another is to combine an iterative back-end for loop closure. Both framework provides reliable performance and have been applied to real applications. Dominated by the fast Fourier transform (FFT), our non-iterative front-end is only of complexity $O(n \lg n)$, where n is the number of points. The map fusion is conducted by element-wise operation, so that both time and space complexity are further reduced. Extensive experiments show that, due to the lightweight of the proposed front-end, the two frameworks of NI-SLAM can run at a much faster speed and yet still with comparable accuracy with the state-of-the-arts.

Convolutional neural networks (ConvNets) have enabled the state-of-the-art performance in computer vision and machine learning. However, little efforts have been devoted to establishing convolution in non-linear space. A new operation, kervolution (kernel convolution), is defined to approximate the non-linear behavior of the human perception system. It generalizes traditional convolution and increases the model capacity without introducing more parameters. Similar to convolution, kervolution can be calculated as element-wise multiplication via Fourier transform. The networks consisting of kervolutional layers are named as kervolutional neural networks (KervNets). The extensive experiments show that KervNets achieve better performance than ConvNets on CIFAR and ImageNet dataset.

In summary, the dissertation demonstrates the superiority of the proposed kernel tools, including KCC, CF, NI-SLAM and KervNets for several visual perception problems. With the kernel tools, we may expect their usage in more applications.

Publication List

[1]. **Chen Wang**, Junsong Yuan, Lihua Xie, “Non-Iterative SLAM," International Conference on Advanced Robotics (ICAR), pp. 83–90, July 2017.

[2]. **Chen Wang**, Handuo Zhang, Thien-Minh Nguyen, Lihua Xie, “Ultra-Wideband Aided Fast Localization and Mapping System," 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2017), pp. 1602-1609, Sep. 2017.

[3]. **Chen Wang**, Le Zhang, Lihua Xie, Junsong Yuan, “Kernel Cross-Correlator," The Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18), Feb. 2018.

[4]. **Chen Wang**, Tete Ji, Thien-Minh Nguyen, Lihua Xie, “Correlation Flow: Robust Optical Flow Using Kernel Cross-Correlators," International Conference on Robotics and Automation (ICRA), 2018.

[5]. **Chen Wang**, Lihua Xie, Junsong Yuan, “Non-Iterative Localization and Dense Mapping," arXiv:1710.05502.

[6]. **Chen Wang**, Xu Fang, Thien-Minh Nguyen, Lihua Xie, “Graph Optimization Approach to Localization with IMU and Ultra-Wideband Measurements," arXiv:1802.10276.

[7]. **Chen Wang**, Jianfei Yang, Lihua Xie, and Junsong Yuan, “Kervolutional Neural Networks," in review.

Patent and Technical Disclosure List

[1]. **Chen Wang***, Lihua Xie*, and Junsong Yuan. "Simultaneous Localization and Mapping Methods and Apparatus." International/U.S. Patent No. PCT/SG2017/050446, Sep. 2016. (*Joint first authors)

[2]. **Chen Wang***, Lihua Xie*, Jianfei Yang, and Junsong Yuan. "Kervolutional Neural Networks." Singapore Provisional Patent Application No. 10201802135Y. NTU Ref: PAT/344/17/18/SG PRV, Mar. 2018. (*Joint first authors)

[3]. **Chen Wang***, Lihua Xie*, and Xu Fang. " A Novel Graph Optimization Approach to Localization with Application for Unmanned Systems.". NTU Ref: TD/254/17, Sep. 2017. (*Joint first authors)

Award List:

[1]. Best Paper Award in robotic planning, “Non-Iterative SLAM” in 2017 18th International Conference on Advanced Robotics (ICAR).

[2]. Championship Award 1st, Category D, Semi-Autonomous, 2015. in Singapore Amazing Flying Machine Competition (SAFMC)

THESIS MIND MAP

Chen Wang | March 22, 2018