A 203 Fps VLSI Architecture Of Improved Dense Trajectories For Real-Time Human Action Recognition

{ Zhi-Yi Lin, Jia-Lin Chen, Liang-Gee Chen } @ National Taiwan University

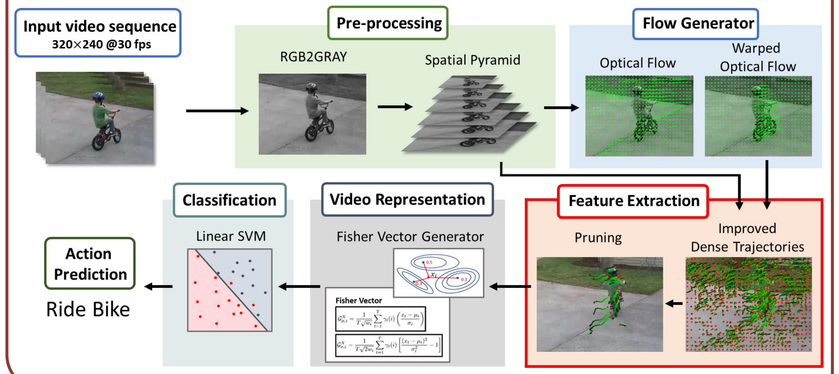
Problem

Few methods can really achieve real-time performance with CPUs. Hardware acceleration is a must in real applications of action recognition.

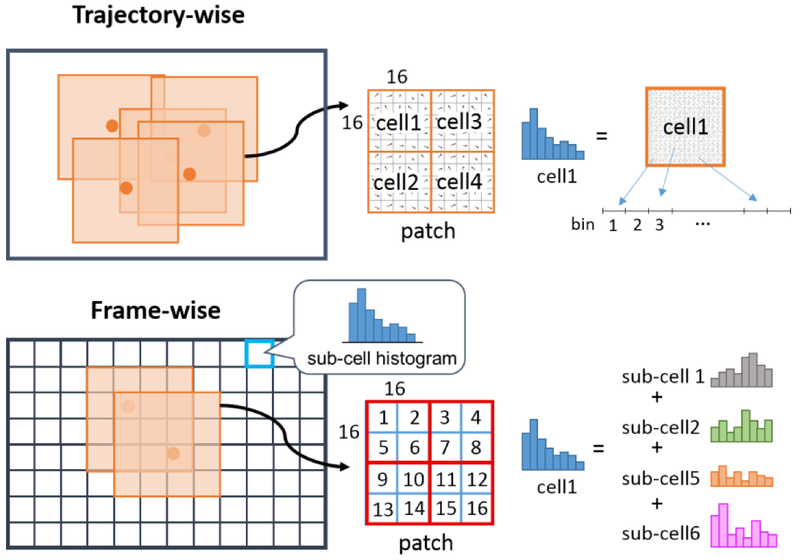
Contributions

This work is the first to propose a hardware-friendly algorithm and an ASIC implementation of Improved Dense Trajectories for real-time action recognition. With our high throughput, low on-chip memory and computation flexibility, our chip can definitely be applied to many real-time applications on mobile devices or be combined with deep learning engines to implement state-of-the-art action recognition systems.

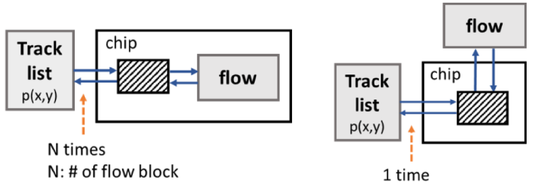
Algorithm Overview



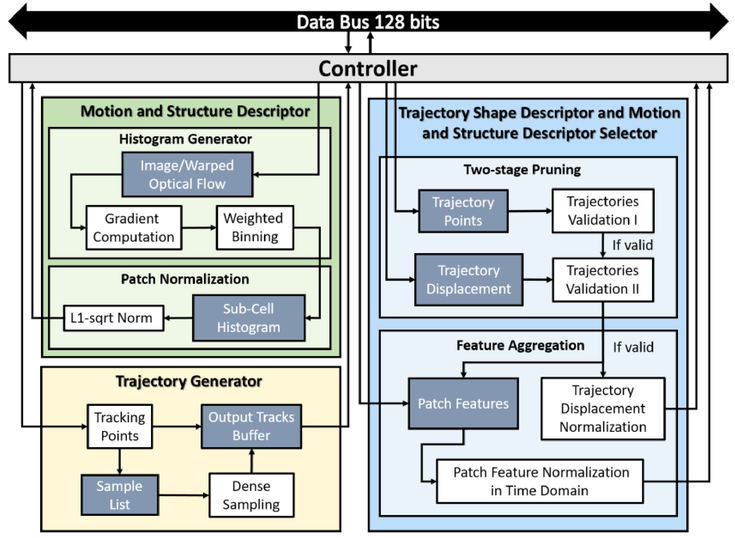
Hardware-Oriented Algorithm



	Architecture 1	Architecture 2
Worst case (46,080 pts/frame)	10,971 + 18,432 × N	46,080 + 18,432 = 64512
Normal case (8600 pts/frame)	10,971 + 1720 × 2 × N =17581 (N=2)	8600 + 1720 × 2 =12,040



Proposed Architecture



ASIC Implementation Result

	iDT
Technology	TSMC 40nm
Chip Area	1.76 mm × 1.76 mm = 3.1 mm^2
Frequency	200 MHz
Gate Count	476.42k
Memory	40.8 kB
Bus Width	128 bits
Throughput (fps)	127 / 203 *
Throughput (track/s)	390k / 624k *
Bandwidth	2.4 GB/s
Power Consumption	58.44 mW