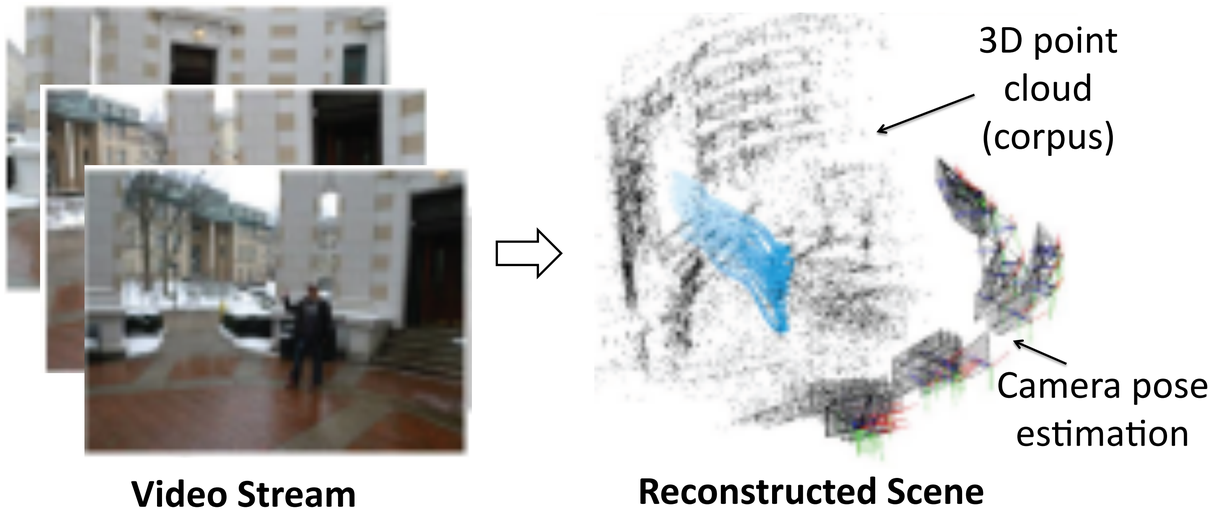
REALTIME 3D RECONSTRUCTION OF REALWORLD SCENES

Yu Wang, Hyun Soo Park, Peter Milder, James C. Hoe, Yaser Sheikh (CMU); Eriko Nurvitadhi, Mei Chen (Intel)

(1) OVERVIEW



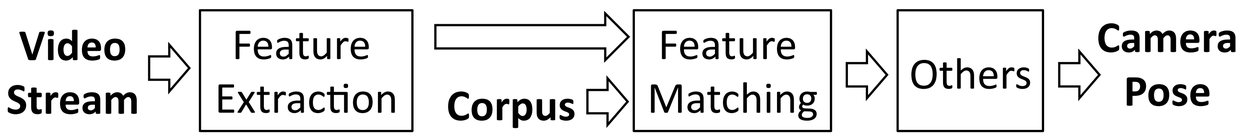
Problem: Perceptually responsive systems require realtime 3D knowledge of environment

Example: to track user movements in a space, and the nearby objects in that space that user can interact with

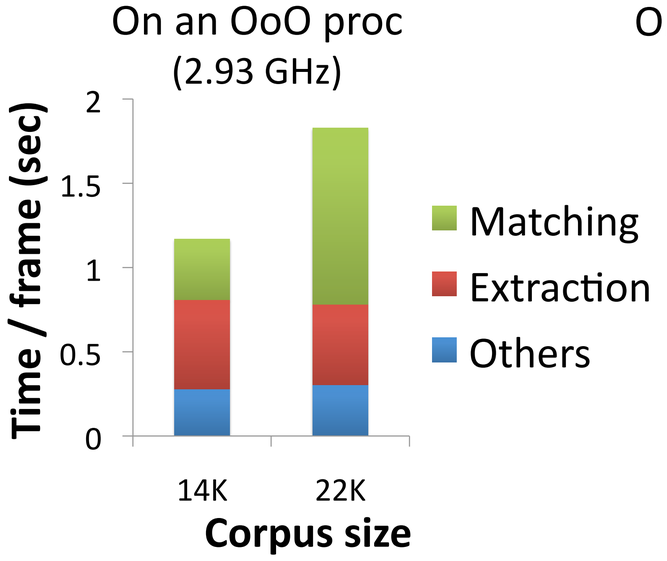
Project goal: produce a low energy, memory-bandwidth efficient embedded solution for realtime 3D reconstruction of dynamic environments from monocular video

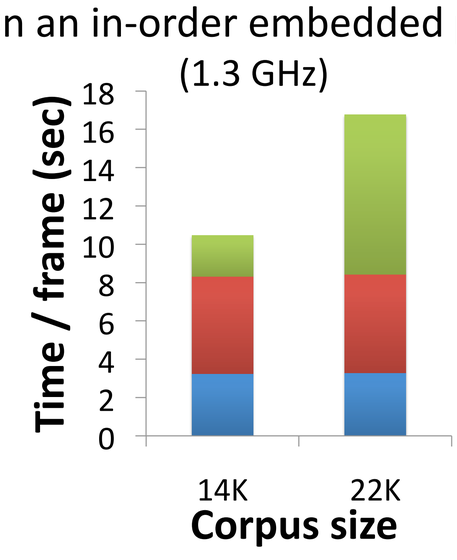
(2) CAMERA POSE ESTIMATION

Given a video stream from a camera, and a 3D point cloud corpus of the area, estimate camera pose w.r.t the corpus



Performance of software (SW) implementation





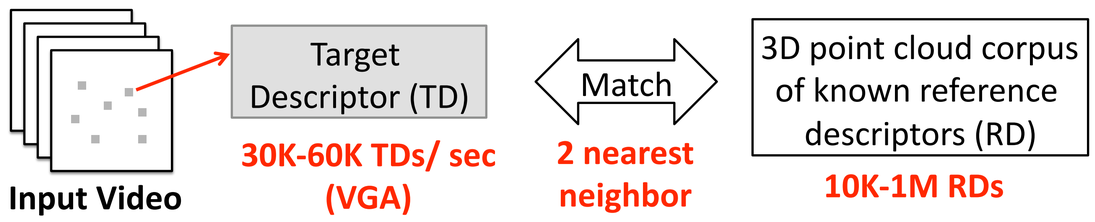
Note: we use ***SIFT*** feature, and our SW uses ***fast approximate*** method for matching. The accuracy of this method may become ***unacceptable*** with larger corpus sizes. ***Exact*** matching is preferred, but very slow to do in SW (for our experiments, ~10x slower than approximate)

SW is too slow even on an aggressive OoO proc (1+ sec/frame), and much worse on an embedded proc

As corpus size increases with larger and/or more complex scenes, feature matching becomes the bottleneck

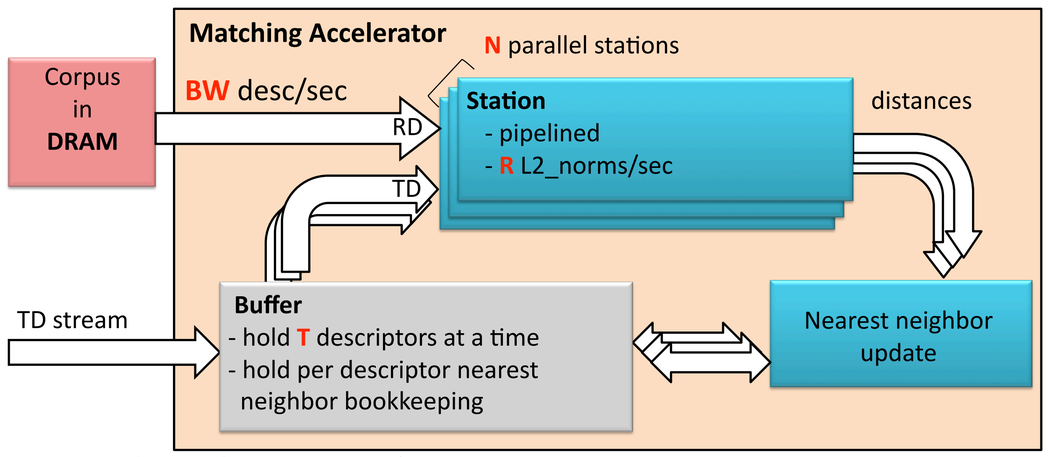
(3) FEATURE MATCHING ACCELERATOR

Feature matching: for each feature in each incoming image, find 2 nearest neighbor (L2-norm) in 3D point cloud corpus



We offload the feature matching bottleneck onto HW

- custom HW architecture to do ***exact*** matching
- orders of magnitude more efficient than SW implementation



Compute-bound: R·N/corpus_size TD per sec
Memory-bound : T·BW/corpus_size TD per sec

(4) FPGA PROOF-OF-CONCEPT

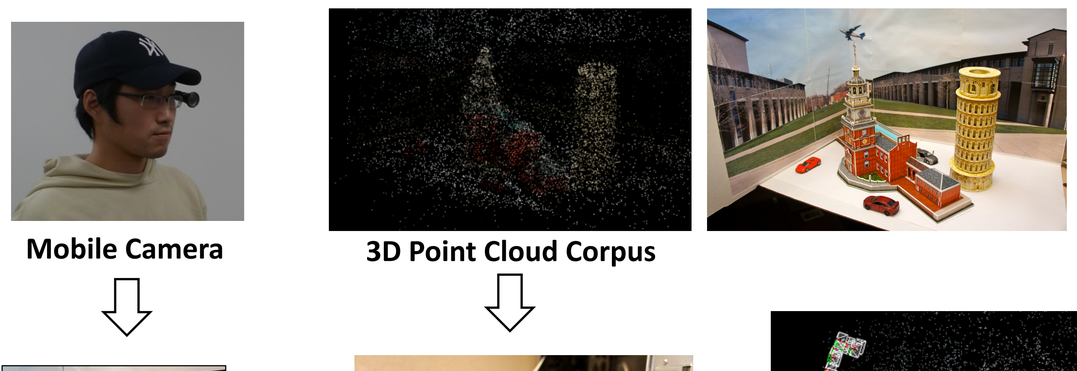
Feature matching on Altera DE4 board (Stratix IV GX FPGA)

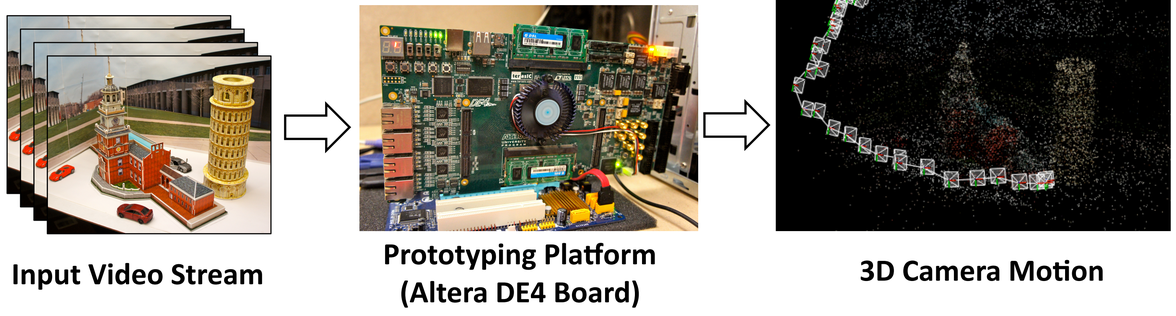
- utilizes 10 parallel stations, runs at 150 MHz clock freq
- occupies ~200K (50%) FPGA logic & 1MB (47%) on-chip mem
- 15 ms & 32 ms per VGA frame (14K & 22K corpus)

24x & 144x faster over SW on OoO proc (14K & 22K corpus)

32x & 257x faster over SW on embedded proc (14K & 22K corpus)

Realtime speed (31+ fps)





What's next

- Accelerate feature extraction
- 3D point cloud corpus reconstruction in realtime