D-PET: A Direct 6 DoF Pose Estimation and Tracking System on Graphics Processing Units

Hung-Yu Tseng, Po-Chen Wu, Yu-Sheng Lin, Shao-Yi Chien
Media IC & System Lab, Graduate Institute of Electronics Engineering, National Taiwan University

**Introduction**

Real-time recovering 6 degrees of freedom (DoF) object pose is essential in augmented reality (AR) and robotics applications. We built a pose estimation and tracking system that:
- Based on [1] and the proposed 3-scale pose search
- Achieves real-time computation on NVIDIA Jetson TX1
- Able to obtain accurate poses of arbitrary types of target.

**Designs**

- The system aims to find the pose with minimum appearance distance
  \[ E_a(p) = \frac{1}{n_t} \sum_{i=1}^{n_t} |l_c(u_i) - l_t(x_i)| \]
  \[ x = (X, Y) \]
  \[ u = (x_c, y_c) \]  
- Pose Estimation Unit
  - Computes the initial pose
  - Implemented based on the approximated pose estimation (APE) scheme
  - Processes each of the poses per pixel
  - Special memory allocation accelerates the parallel computation
    - 4-2 floats memory allocation for storing pose parameters
    - Camera image stored in texture memory
    - Accesses target information via load uniform instructions (LDU)
- Pose Tracker
  - Realizes accurate pose tracking
  - Applies a 6D pose search pattern to perform 3-scale pose search
    - 5 points in each of the rotation dimensions, 7 points in each of the translation dimensions, totally 42875 points for the search pattern

**Results**

- Evaluation
  - Rotation error (degree): \( E_R = \text{acosd} \left( \frac{\text{Tr}(R_1 R_2^{-1})}{\sqrt{\text{det}(R_1 R_2)}} \right) \)
  - Translation error (%): \( E_t = \frac{\| t - t_f \|}{\| t_f \|} \times 100 \)
- Compare Pose Estimation Unit with APE [1] and ASIFT using the synthetic dataset [1]
- Evaluate Pose Tracker using the real dataset [2]

**Conclusion**

We propose D-PET, a direct 6 DoF Pose Estimation and Tracking system implemented on an embedded GPU. Compared to the state-of-the-art feature based systems, it is able to deal with more general targets and performs favorably in terms of accuracy and robustness.

Reference:


https://goo.gl/9CCfcV