

# Automatic Pose Estimation for Range Images on the GPU

Marcel Germann\*, Michael D. Breitenstein\*, In Kyu Park†‡, Hanspeter Pfister†‡

- \* Swiss Federal Institute of Technology Zurich, Switzerland
- †‡ Inha University, Incheon, Korea
- ‡ Mitsubishi Electric Research Laboratories (MERL), Cambridge, USA

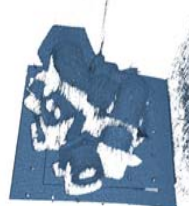
## Abstract

Object pose (location and orientation) estimation is a common task in many computer vision applications. Although many methods exist, most algorithms need manual initialization and lack robustness to illumination variation, appearance change, and partial occlusions.

We propose a fast method for automatic pose estimation without manual initialization based on shape matching of a 3D model to a range image of the scene. We developed a new error function (see the paper) to compare the input range image to pre-computed range maps of the 3D model. We use the tremendous data-parallel processing performance of modern graphics hardware to evaluate and minimize the error function on many range images in parallel. Our algorithm is simple and accurately estimates the pose of partially occluded objects in cluttered scenes in about half second.

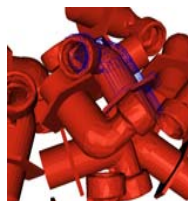
## Goal

- Fast and robust pose estimation
- Given: 3D model
- Input: depth image (e.g. from laser scanner)

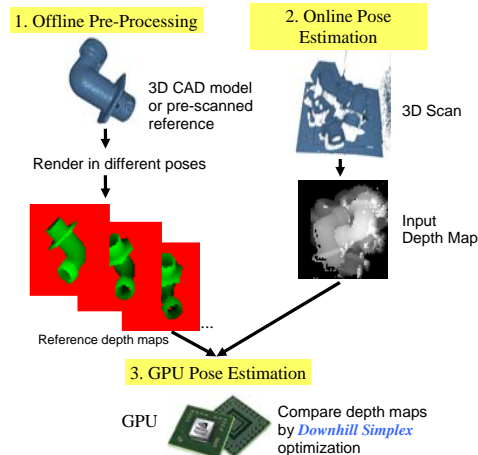


## Application: Robot Bin Picking

Industrial robot selects and picks top-most object from a bin of identical or different objects



## Overview of the Proposed Approach



- Simple algorithm but massively parallelized on the GPU
- Alignment by parallelized testing of many pose hypotheses (e.g. 2,048)

### Improvement using NVIDIA CUDA (Compute Unified Device Architecture)

- Flexible scatter operation
- Linear memory within whole VRAM space
  - More flexibility in data structure
  - Remove the background pixel of reference view
- Use of very fast cache memory (shared memory)
- (+) No dependence on OpenGL and GLSL
- (-) Dependence on NVIDIA GPU

## Results and Demo with CUDA

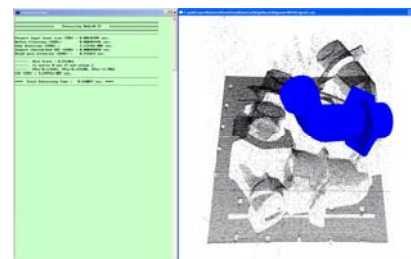
Comparison of running time (in sec) with 2,048 reference views

Processor	Median	Edge	EDT	Pose	Total	Gain
GPU (GLSL and FBO on G80)	0.0007	0.0003	0.0018	1.0247	1.0278	30.9781
CPU (Intel Pentium D 945)	0.0008	0.0002	0.0141	31.8236	31.8393	-
GPU (CUDA on G80)	0.0003	0.0004	0.0015	1.3016	1.3038	24.4207
GPU (Optimized CUDA on G80)	0.0003	0.0002	0.0013	0.5167	0.5186	61.3921

(Result in the paper)

(New result)

Real time demo 1 (real scene)



Real time demo 2 (virtual bin-picking system)

