Best view estimation for active exploration and object recognition

ProDEI - Advanced Methods of Modeling and Simulation

Carlos M. Costa
carlos.costa@fe.up.pt

Faculty of Engineering
University of Porto
Portugal

February, 2017

Presentation Outline

CONTEXT

- 3D object recognition is a challenging task that may require active perception of the environment
- Simulating sensor 3D depth data from a set of representative environments can help decide the type, number and disposition of sensors that maximize the observable target objects surface area within a given work area

RESEARCH AREAS

- 3D modeling and simulation of
  - The 3D environment, including physics
  - The different types of depth sensors
- 3D rendering with occlusions
- Active perception
- Object recognition
- Bin picking

SOFTWARE DEPENDENCIES

- Robot Operating System (ROS)
  - For fast integration between simulation and real robots and 3D visual inspection (Rviz)
- Gazebo simulator
  - For world / sensor simulation and 3D rendering
- Point Cloud Library (PCL)
  - For point cloud processing

Fig. 1: Main software dependencies

DEPTH SENSORS MODELING

- Modeling of 8 different types of depth sensors
  - Structured light sensors
    - Asus Xtion Pro Live
    - Ensenso N35
    - Intel RealSense SR300
    - Kinect XBox 360
    - Kinect XBox One
    - Orbbec Astra
  - Stereo sensors
    - MultiSense S7
    - ZED stereo camera
- Pin-hole camera model coupled with OpenGL color rendering and depth buffer allows to specify the sensors:
  - Resolution (pixels)
  - Field of view (radians)
  - Minimum and maximum measurement range (meters)
  - Sensor acquisition rate (Hz)

DEPTH SENSORS MODELING

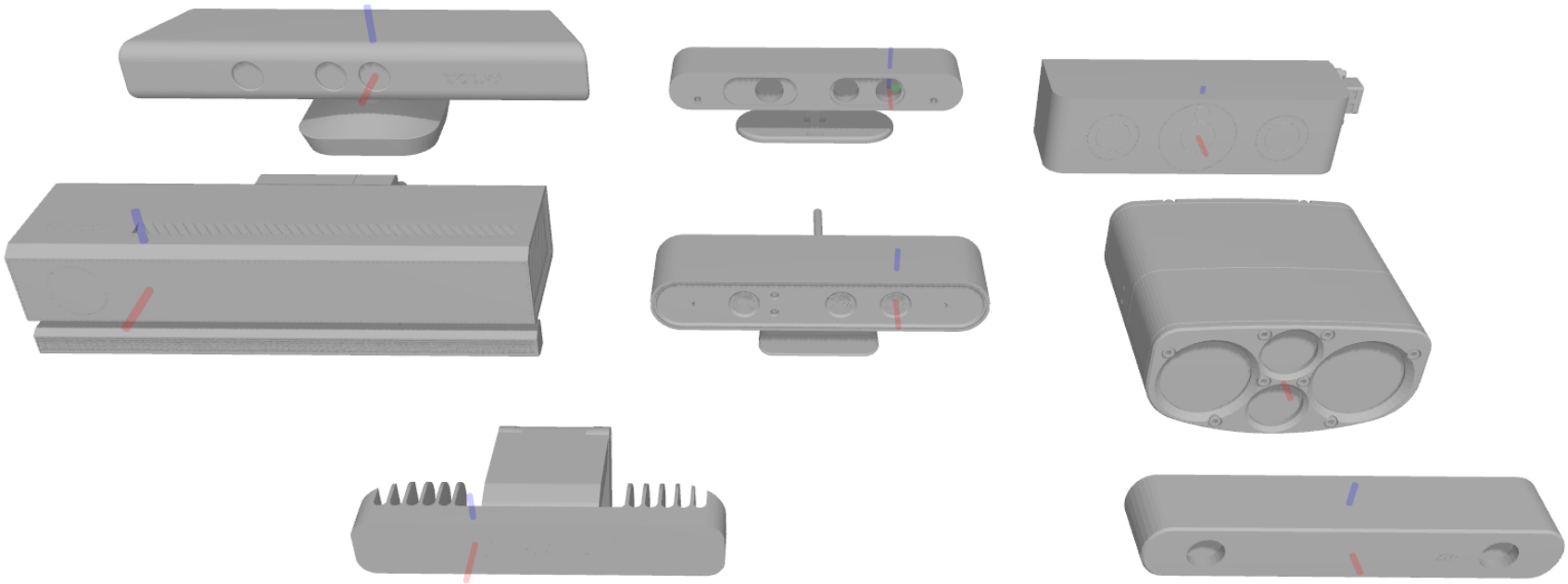

Fig. 2: Sensors 3D CAD models with the display of the depth image coordinate frames (horizontally from top left to bottom right: Kinect XBox 360, Asus Xtion Pro Live, Ensenso N35, Kinect XBox One, Orbbec Astra, MultiSense S7, Intel RealSense SR300, ZED stereo camera)

ENVIRONMENT MODELING

- Modeling of 4 different test environments
  - 1 for active perception, with hand occlusions
  - 1 for bin picking with minimal occlusions
  - 1 for bin picking with significant occlusions
  - 1 for multiple object bin picking with several occlusions
- Modeling of 3 bin picking objects
  - Starter motor (target object for bin picking)
  - Alternator (for bin picking occlusions)
  - Differential gearbox (for bin picking occlusions)
- Modeling of 2 support objects
  - Trolley with shelves
  - Large stacking box
- Target objects use a special surface material that ignores light effects (such as light shading, shadows) and have a specific color (pure green) that will be used for sensor data segmentation. It was used .stl CAD models for 3D rendering and .ply point clouds for 3D data processing

ACTIVE PERCEPTION ENVIRONMENT

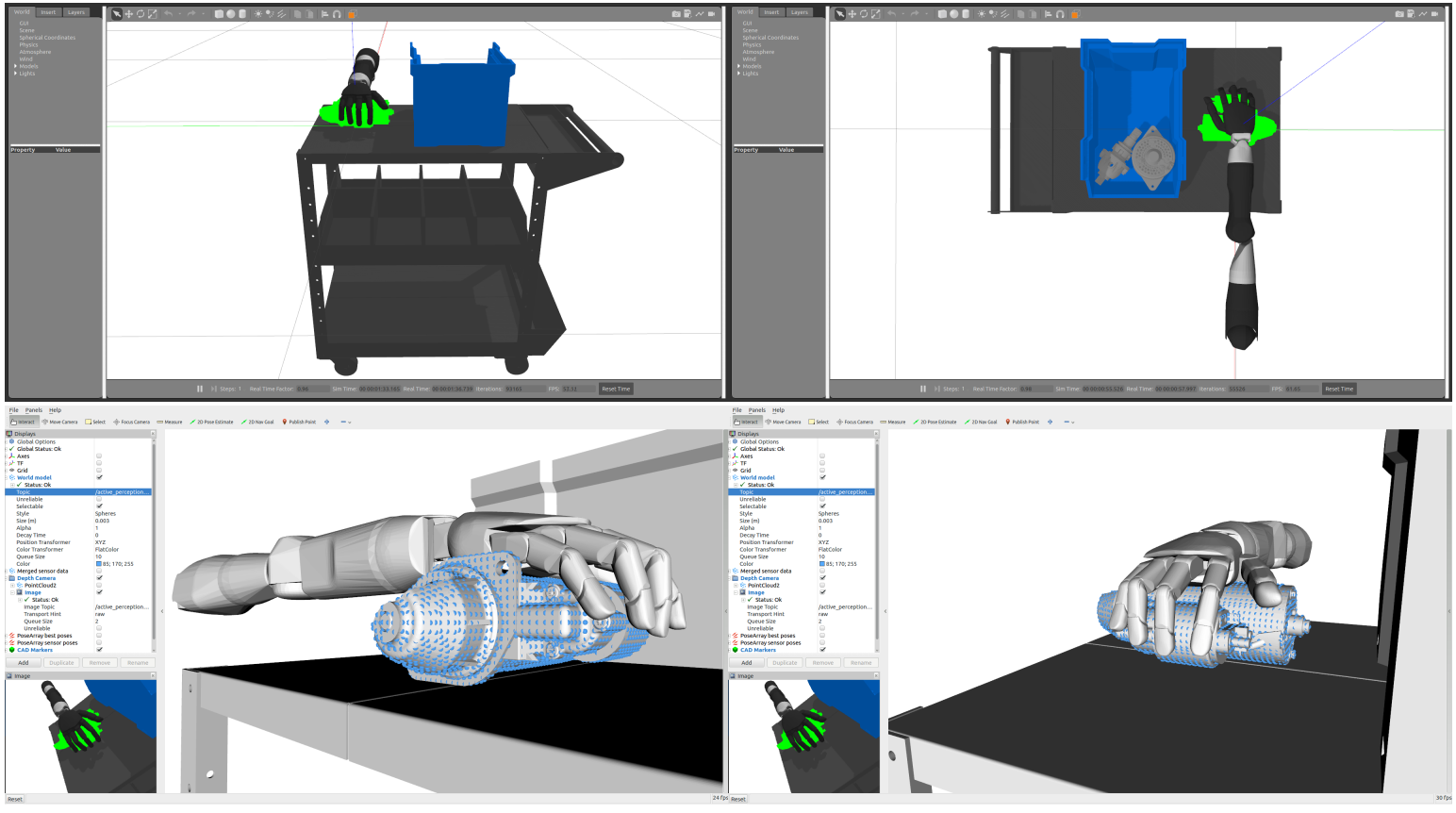

Fig. 3: Active perception environment renderings from Gazebo with target objects in green (top images) and associated CAD model point clouds displayed as blue spheres in Rviz (bottom images)

BIN PICKING ENVIRONMENT

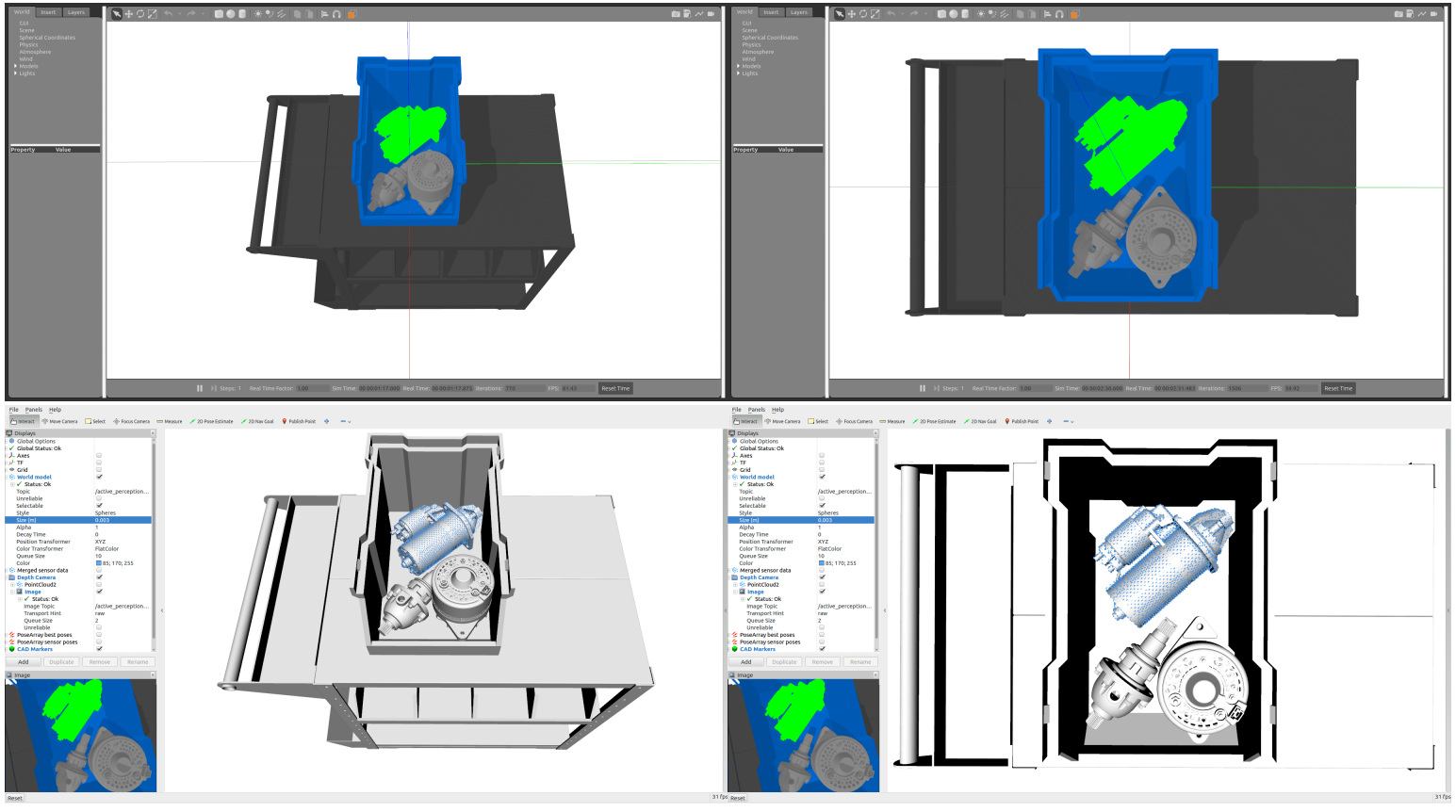

Fig. 4: Bin picking environment renderings from Gazebo with target objects in green (top images) and associated CAD model point clouds displayed as blue spheres in Rviz (bottom images)

BIN PICKING WITH OCCLUSIONS ENVIRONMENT

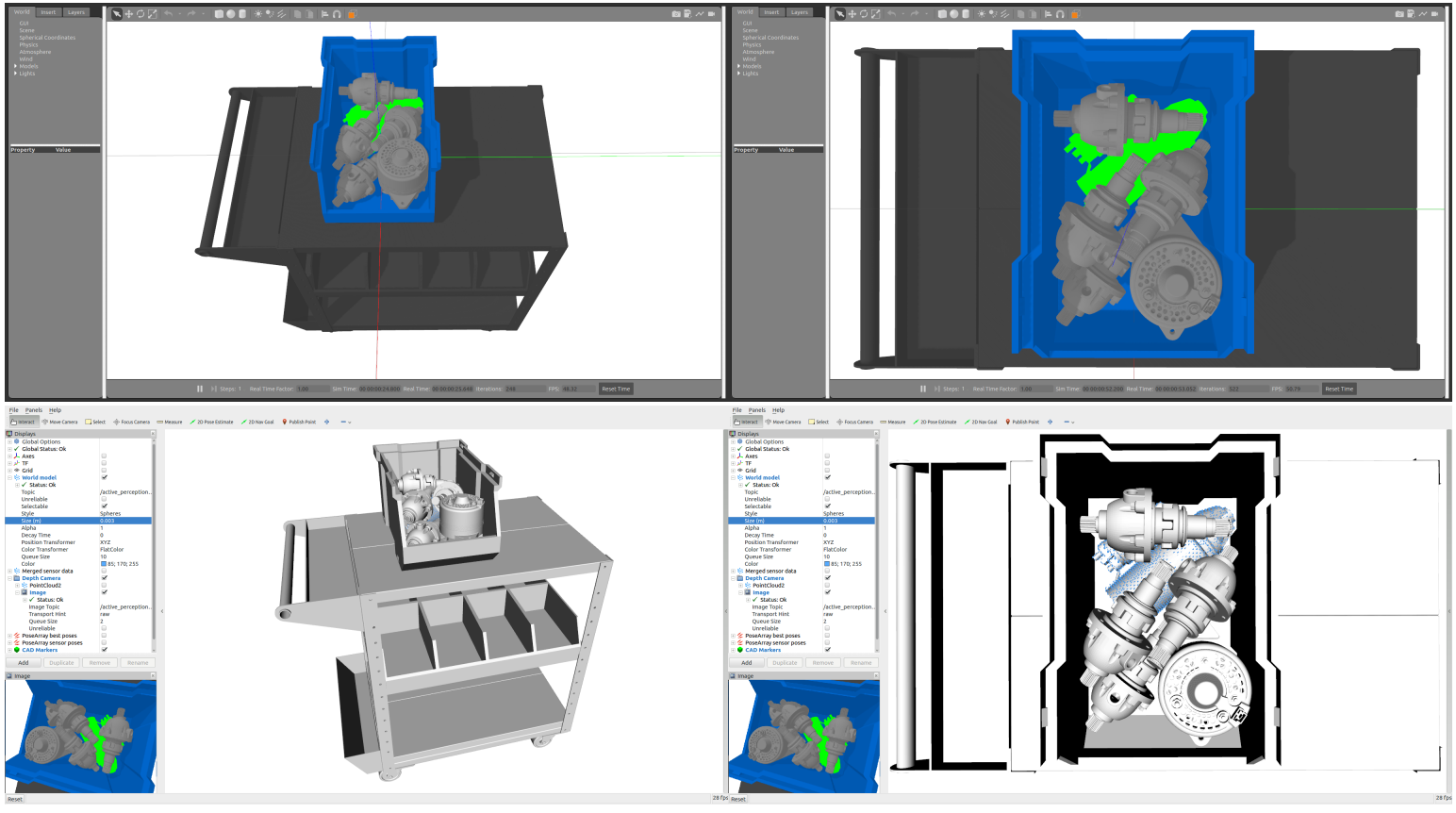

Fig. 5: Bin picking with occlusions environment renderings from Gazebo with target objects in green (top images) and associated CAD model point clouds displayed as blue spheres in Rviz (bottom images)

MULTIPLE BIN PICKING WITH OCCLUSIONS ENVIRONMENT

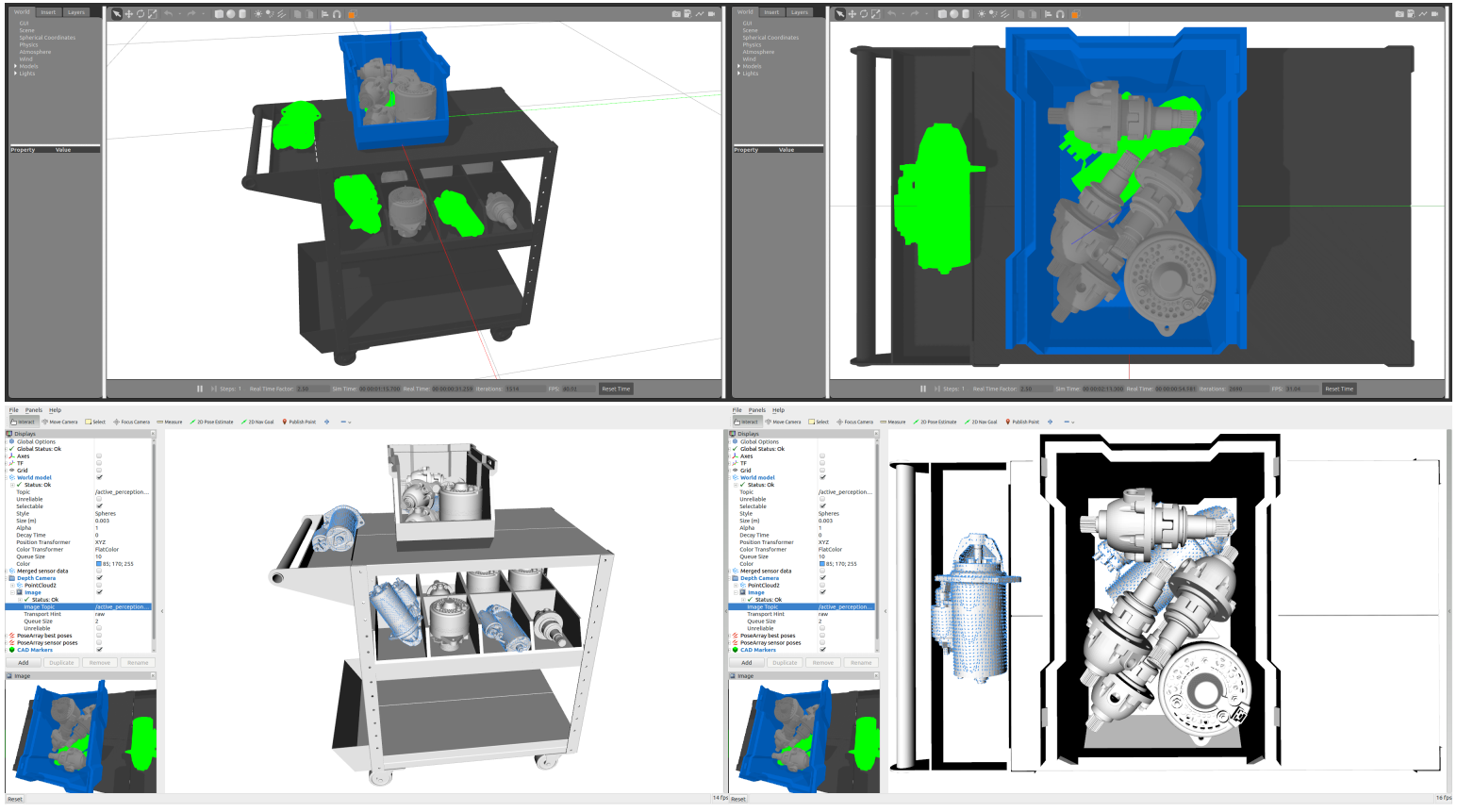

Fig. 6: Multiple bin picking with occlusions environment renderings from Gazebo with target objects in green (top images) and associated CAD model point clouds displayed as blue spheres in Rviz (bottom images)

ESTIMATION OF THE BEST SENSOR CONFIGURATION

- Given a set of observation positions
  - If only one sensor is available
    - Analyze all sensors data from each observation position
    - Apply a voxel grid space partition to avoid reduce variability of the estimated surface area observed by each sensor
    - Choose the sensor that can observe the most surface area of the objects

ESTIMATION OF THE BEST SENSOR CONFIGURATION

- Given a set of observation positions
  - If several sensors area available (type and quantity adjustable)
    - Extract point clouds for each observation point
    - Using a RANSAC approach, choose randomly a set a N sensors
    - Merge the sensor data and estimate the observable surface area
    - At the end of a given number of iterations or if the observable area reaches a given minimum, terminate the search, keeping the best sensor configuration found

Thank you!
Questions?