Perception and Planning for Humanoid Disaster-response Robots

Sven Behnke
University of Bonn
Computer Science Institute VI
Autonomous Intelligent Systems
Some of our Cognitive Robots

- Equipped with numerous sensors and actuators
- Complex demonstration scenarios

Soccer  Domestic service  Mobile manipulation  Bin picking  Aerial inspection
Motivation

■ Capabilities of disaster-response robots were insufficient for providing effective support to rescue workers.
  ● Mobility: difficulties with uneven terrain, stairs, and debris
  ● Manipulation: only a single actuator with simple end-effectors
  ● User interface: requires extensive training, not intuitive, situation awareness problematic

■ Complexity of achievable tasks and execution speed are low
Mobile Manipulation Robot Momaro

- Four compliant legs ending in pairs of steerable wheels
- Anthropomorphic upper body
- Sensor head
  - 3D laser scanner
  - IMU, cameras

[Schwarz et al. Journal of Field Robotics 2017]
Manipulation Operator Interface

- 3D head-mounted display
- 3D environment model + images
- 6D magnetic tracker

[Rodehutskors et al., Humanoids 2015]
Local Multiresolution Surfel Map

- Registration and aggregation of 3D laser scans
- Local multi-resolution grid
- Surfel in grid cells

[Droeschel et al., Robotics and Autonomous Systems 2017]
Filtering Dynamic Objects

- Maintain occupancy in each cell
- Remove measurements of empty cells
Allocentric 3D Mapping

- Registration of egocentric maps by graph optimization

[Droeschel et al., Robotics and Autonomous Systems 2017]
Valve Turning Interface

- Align wheel model with 3D points using interactive marker

[Schwarz et al. Journal of Field Robotics 2017]
Debris Tasks
Team NimbRo Rescue

Best European Team (4th place overall), solved seven of eight tasks in 34 minutes
DLR SpaceBot Cup 2015

- Mobile manipulation in rough terrain

[Schwarz et al., Frontiers on Robotics and AI 2016]
Autonomous Mission Execution

- 3D mapping, localization, mission and navigation planning

- 3D object perception and grasping

[Schwarz et al. Frontiers 2016]
Navigation Planning

- Costs from local height differences
- A* path planning

[Schwarz et al., Frontiers in Robotics and AI 2016]
3D Map
Improved Sensor Head

- Continuously rotating Velodyne Puck VLP-16
  - 300,000 3D points/s
  - 100 m range
  - Spherical field of view
- Three wide-angle color cameras (total FoV 210×103°)
- Kinect V2 RGB-D camera on pan-tilt unit
3D Map of Indoor+Outdoor Scene

[Droeschel et al., Robotics and Autonomous Systems 2017]
Considering Robot Footprint

- Costs for individual wheel pairs from height differences
- Base costs
- Non-linear combination yields 3D \((x, y, \theta)\) cost map

[Klamt and Behnke, IROS 2017]
3D Driving Planning \((x, y, \theta)\): A* 

- 16 driving directions

- Orientation changes

=> Obstacle between wheels

[Klamt and Behnke, IROS 2017]
Making Steps

- If not drivable obstacle in front of a wheel
- Step landing must be drivable
- Support leg positions must be drivable

[Klamt and Behnke: IROS 2017]
Planning for Challenging Scenarios

[Klamt and Behnke: IROS 2017]
Centauro Robot

- Serial elastic actuators
- 42 main DoFs
- Schunk hand
- 3D laser
- RGB-D camera
- Color cameras
- Two GPU PCs

[Tsagarakis et al., IIT 2017]
Main Operator Telepresence Interface

- Tendon-driven dual-arm exoskeleton
- Active wrist with differential tendon transmission
- Underactuated hand exoskeleton
- Head-mounted display
- Foot pedals

[Frisoli et al., SSSA 2017]
Main Operator Control

Manipulation Tasks
- Surface
- Valve (lever)
- Valve (gate)
- Snap hook
- Fire hose
- 230V connector
- Cutting tool
- Driller
- Screw driver
- Grasping

Used control interfaces
- Joystick
- Exus
- 6D
- Keyframes
- Stepping
- Autonomous
Turning a Valve
Connecting a Plug

Manipulation Tasks
- Surface
- Valve (lever)
- Valve (gate)
- Snap hook
- Fire hose
- 230V connector
- Cutting tool
- Driller
- Screw driver
- Grasping

Used control interfaces
- Joystick
- Exus
- 6D
- Keyframes
- Stepping
- Autonomous
3D Mapping and Localization
Walking over a Step Field
Terrain Classification

Schilling et al., IROS 2017

Color image → Augmented & fine-tuned CNN → Visual features → Random Forest → Traversability class

- safe (green)
- risky (yellow)
- obstacle (red)

Registered point cloud → Feature extraction (Height, Slope, Roughness) → Geometric features

[Schilling et al., IROS 2017]
## Hybrid Driving-Stepping Locomotion Planning: Abstraction

<table>
<thead>
<tr>
<th>Level</th>
<th>Map Resolution</th>
<th>Map Features</th>
<th>Robot Representation</th>
<th>Action Semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.5 cm 64 orient.</td>
<td>Height</td>
<td></td>
<td>Individual Foot Actions</td>
</tr>
<tr>
<td>2</td>
<td>5.0 cm 32 orient.</td>
<td>Height, Height Difference</td>
<td></td>
<td>Foot Pair Actions</td>
</tr>
<tr>
<td>3</td>
<td>10 cm 16 orient.</td>
<td>Height, Height Difference, Terrain Class</td>
<td></td>
<td>Whole Robot Actions</td>
</tr>
</tbody>
</table>

[Klamt and Behnke, IROS 2017, ICRA 2018]
Deep Learning Object Detection

[Johnson et al. 2015]
CENTAUNRO Workspace Perception Data Set

129 frames, 6 object classes

https://www.centauro-project.eu/data_multimedia/tools_data
Tool Detection Results

[Schwarz et al. IJRR 2017]
Tools Detection Examples

[Schwarz et al. IJRR 2017]
Semantic Segmentation

- Deep CNN

Pixel-wise accuracy:

<table>
<thead>
<tr>
<th></th>
<th>Clamp</th>
<th>Door handle</th>
<th>Driller</th>
<th>Extension</th>
<th>Stapler</th>
<th>Wrench</th>
<th>Background</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.727</td>
<td>0.751</td>
<td>0.769</td>
<td>0.889</td>
<td>0.775</td>
<td>0.734</td>
<td>0.992</td>
<td>0.805</td>
</tr>
</tbody>
</table>
RefineNet for Semantic Segmentation

- Scene represented as feature hierarchy
- Course-to-fine semantic segmentation
- Combine higher-level features with missing details

[Lin et al. CVPR 2017]
The Data Problem

- Deep Learning in robotics (still) suffers from shortage of available examples
- We address this problem in two ways:

1. **Generating data:**
   Automatic data capture, online mesh databases, scene synthesis

2. **Improving generalization:**
   Object-centered models, deformable registration, transfer learning, semi-supervised learning
Object Capture and Scene Rendering

- Turntable + DLSR camera

- Rendered scenes

[Schwarz et al. ICRA 2018]
Semantic Segmentation Example

[Schwarz et al. ICRA 2018]
Object Pose Estimation

- Cut out individual segments
- Use upper layer of RefineNet as input
- Predict pose coordinates

[Schwarz et al. ICRA 2018, Periyasamy et al. IROS 2018]
From Turntable Captures to Textured Meshes

Fused & textured result
Transfer of Manipulation Skills

- Objects belonging to the same category can be handled in a very similar manner.
Transfer of Manipulation Skills

Knowledge Transfer
Learning a Latent Shape Space

- Non-rigid registration of instances and canonical model
- Principal component analysis of deformations
Interpolation in Shape Space

[Rodriguez and Behnke ICRA 2018]
Shape-aware Non-rigid Registration

- Partial view of novel instance
- Deformed canonical model

[Rodriguez and Behnke ICRA 2018]
Shape-aware Registration for Grasp Transfer

- Full point cloud
- Partial view
Collision-aware Motion Generation

Constrained Trajectory Optimization:

- Collision avoidance
- Joint limits
- Time minimization
- Torque optimization

[Pavlichenko et al., IROS 2017]
Grasping an Unknown Power Drill
Fastening a Screw
Bimanual Fastening Task
Bimanual Grasping
Bimanual Drilling
Opening a Door with a Key
Closing a Shackle
Bimanual Plug Tasks
Step Field with Debris
Autonomous Navigation
Conclusions

- Developed capable humanoid robot systems for disaster-response scenarios
- Teleoperation is flexible, but demanding and error-prone
- Autonomy for common navigation and manipulation tasks needed
- Challenges include
  - Capable and affordable robot platforms
  - 4D semantic perception
  - High-dimensional motion planning
- Promising approaches
  - Shared autonomy
  - Structured learning