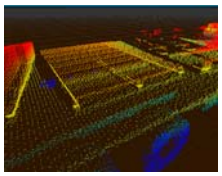







Agile Robotics: Spoken and Gestural Command of Robots in Human Workplaces

RSS Lecture #18

April 14th, 2014

Genesis of Agile Robotics Effort

- Army Logistics Innovation Agency asked us to apply autonomy to military logistics
  - Keep personnel safe; improve efficiency; enable one person to command multiple bots


- Over winter 2007-2008, we developed the concept of a robotic (autonomous) forklift
- DDR&E funded the project in April 2008

Agile Robotics Team

Platform buildup

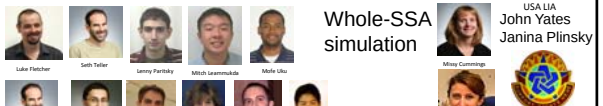
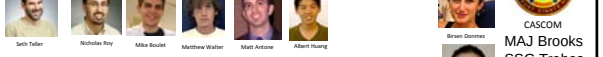
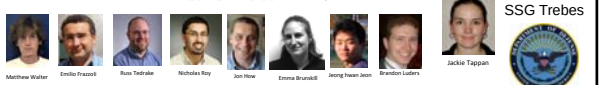
Safety & intent

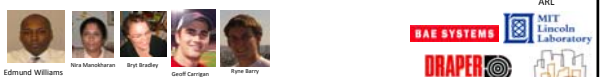
Situational awareness

Planning & control

Supervisor interface

Program support



Whole-SSA simulation

USA LIA
John Yates
Janina Plinsky

CASCOM
MAJ Brooks
SSG Trebes

DDRE

ARL

BAE SYSTEMS

MIT Lincoln Laboratory

DRAPER

State of the Art Is Not Applicable

- Highly prepared & structured environments
  - Indoors; uniform lighting; smooth, densely-barcoded floors; precise prior metrical maps; precision shelving; centralized task planning and path control; no people in workspace!

Kiva Systems


- Army requirements break *every assumption*!

Why Not Just Adapt UAV Systems?

- Require high-precision GPS
  - Good prior maps, live views
- Operate in sparse space (air)
  - Few obstacles, no people
- Assume smooth dynamics
  - Calm air; wind at worst
- Support only simple commands
  - Follow/orbit waypoint; track; strike
- High-latency, distant teleop OK
  - Multiple humans per UAV (relatively few robots)
  - Command station can be fixed at a distant location

Approach Won't Work for UGVs!

- Low-precision, intermittent GPS
  - Poor prior maps; close approach
- Crowded operating space
  - Terrain, obstacles, people
- Fast, complicated dynamics
  - Terrain, collision avoidance, slip/skid
- Must support complex command set
  - Example: following directions in urban area
- Teleop must be low-latency, thus **nearby**
  - Can't use dedicated satellite-based links
  - UGV can't be controlled from moving cabin
    - Humans can't tolerate visual / vestibular disparity
- One human can't teleoperate multiple UGVs
  - Fast dynamics requires human operator's dedicated attention

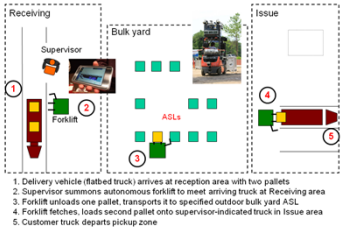
Robotics Rodeo at Fort Hood, 9/09

RIPSAW teleoperated UGV

CAST human-led convoying

AR: Fundamental Design Goals

- Develop unmanned ground robots that can:
  - Operate in outdoor, *unprepared* environments
  - Take direction from people in a *natural* way
  - Work alongside, and be *tolerated* by, people



- Forklift is 1st focus; methods apply broadly

Technical Approach

- Intensive effort to understand military users
- Unprepared environments
  - Reliance on local sensing & decision-making
  - Novel mobile (pallet) manipulation capabilities
- Natural direction by people
  - Speech and gesture interface
- Cultural acceptability to people
  - Multiply redundant safety systems
  - Predictable, human-like behavior
    - Displays internal state, plans to bystanders
  - "Rookie" metaphor: can request, accept help

An "Acceptable" Robot must be:

- **Capable**:
  - …can do some useful task in close quarters
- **Responsive**:
  - …interacts with user, given (limited) commands
- **Safe**:
  - …highly unlikely to hurt people or break things
- **Predictable**:
  - …moves and acts in an unsurprising fashion
- **Subservient**:
  - …(almost) always defers to nearby humans

Platform: Roboticizing a Forklift

Stock, American-made Toyota forklift; manually controlled; weighs 3 tons, lifts 1.5 tons; costs $25K off the lot.

2. Add **sensing**: laser range scanning of surround, of terrain, and of pallets; 360-degree video for supervisory view; IMU + odometry for egomotion; GPS for rough positioning; array microphones to hear shouted warnings and commands

In-House

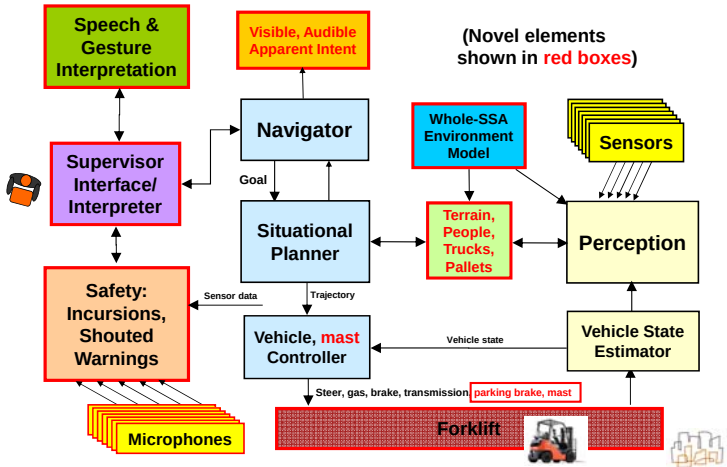
3. Add internal **network** for data routing, **and** five quad-**CPU** laptops for perception, planning, and control; export supervisor views wirelessly

4. Add **supervisor's tablet** with on-board speech and gesture recognition, display of live "robot's-eye" view

5. Add LEDs, signage, and speakers to **annunciate state and intentions** to bystanders

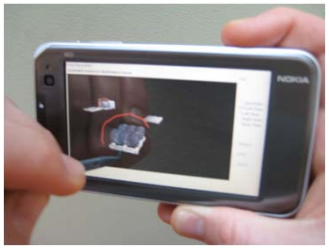
1. Convert to **drive-by-wire**: put gas, brake, steering, parking brake, mast and tine controls under computer control

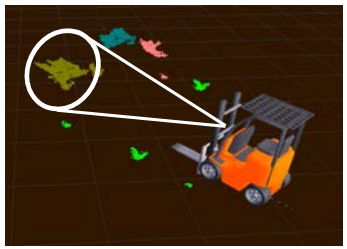
AR System Architecture



Task-Level Speech/Gesture Interface

Speech, Gestures & Target Pallet





- Supervisor gesture and calibrated sensor model used to associate segmented data with correct pallet
- Other gestures indicate placement location
- Supervisor speech recognized in real-time

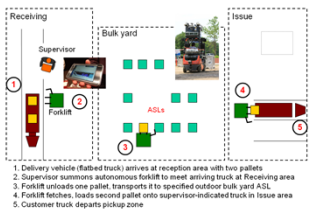
Autonomous Operation (Ft Lee)

Video

Also: remote operation (from AUSA in Washington DC)

Supervisor at AUSA commanding robot 700 km away at MIT over commodity wifi!

Cultural issues (safety, efficiency)



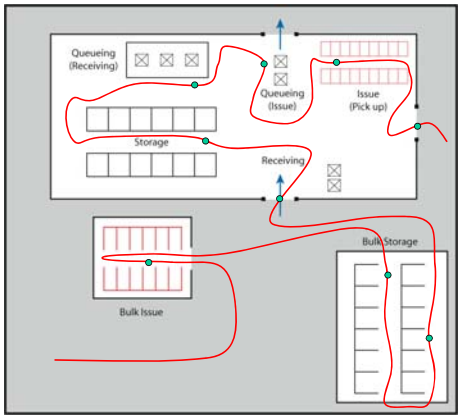
- Warehouse layout, shared knowledge
- Deference to bystanders, other vehicles
- Reasonable, predictable behavior
- Seamless autonomy handoff (rookie model)
- Human presence & shout detection

Semi-Structured Environment

Idea: Give forklift a narrated, guided tour!

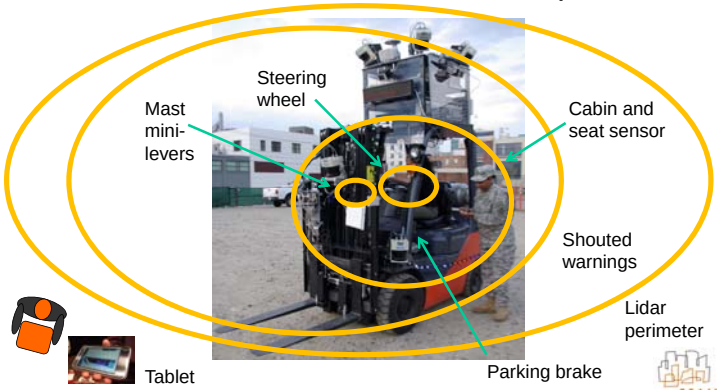
Manually-driven forklift path

Operator utterance and / or gesture



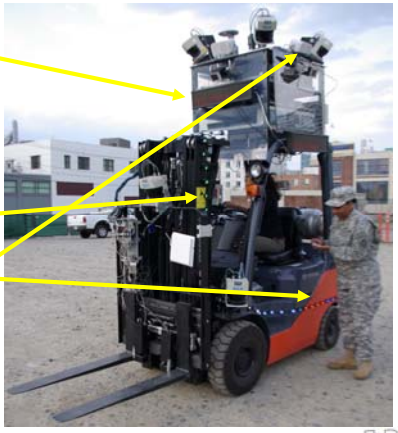
Safer Operation around People

- Perception: six redundant safety layers
- Visible and audible *annunciation* of plans



External Display of State, Plans

- Signage
  - "Stopped: Pedestrian"
- LED strings
  - Imminent motion; bystanders
- Speakers
  - [Bell rings], says "about to move"



Shouted Warning Detection

- Goal: detect shouted "Stop!", "Watch it!", etc.
- Novel aspects:
  - Continuously running speech detection (no push-to-talk)
  - Highly challenging audio environment
  - Humans typically far (several meters) from microphones
- Example data collection from Ft. Campbell:
  - Engine noises, contact with pallet load, motion etc.

What's next: Higher-level Interface

- Single-pallet task interface is tedious to use

Hierarchical Task-Level Autonomy

- Not tele-operated; Not fully autonomous
- "Sliding autonomy" solution in which human supervisor gives step-by-step instructions:
  - Summon forklift to working area
  - Direct it toward one pallet (of several)
  - Help it localize pallet, slots (if necessary)
  - Direct it to destination (bulk lot, issue area etc.)
  - Indicate placement region on ground or flatbed
- Rich, incremental path to full autonomy
  - Ever-shorter utterances, ever-longer tasks
  - Method applicable to most task domains

Conclusions

- Research value of engaging with real users
- Task specificity provides useful structure for technical advances
- Situational awareness as key technical enabler
- Research progress from "cultural" issues

What's next in RSS

- Wednesday:
  - Autor guest lecture on Technology Anxiety; Milestone demos in lab
- Friday
  - No forum; Intermediate grade reports available
- Next Monday (& Tuesday), 21-22 April
  - Patriots Day, no MIT classes or labs
- Next Wednesday, 23 April
  - No lecture; milestone & demo in lab
- Next Thursday, 24 April
  - Drop date
- Next Friday, 25 April
  - No forum
- Monday, 28 April
  - Debates begin! (Here in 32-155, during lecture timeslot)