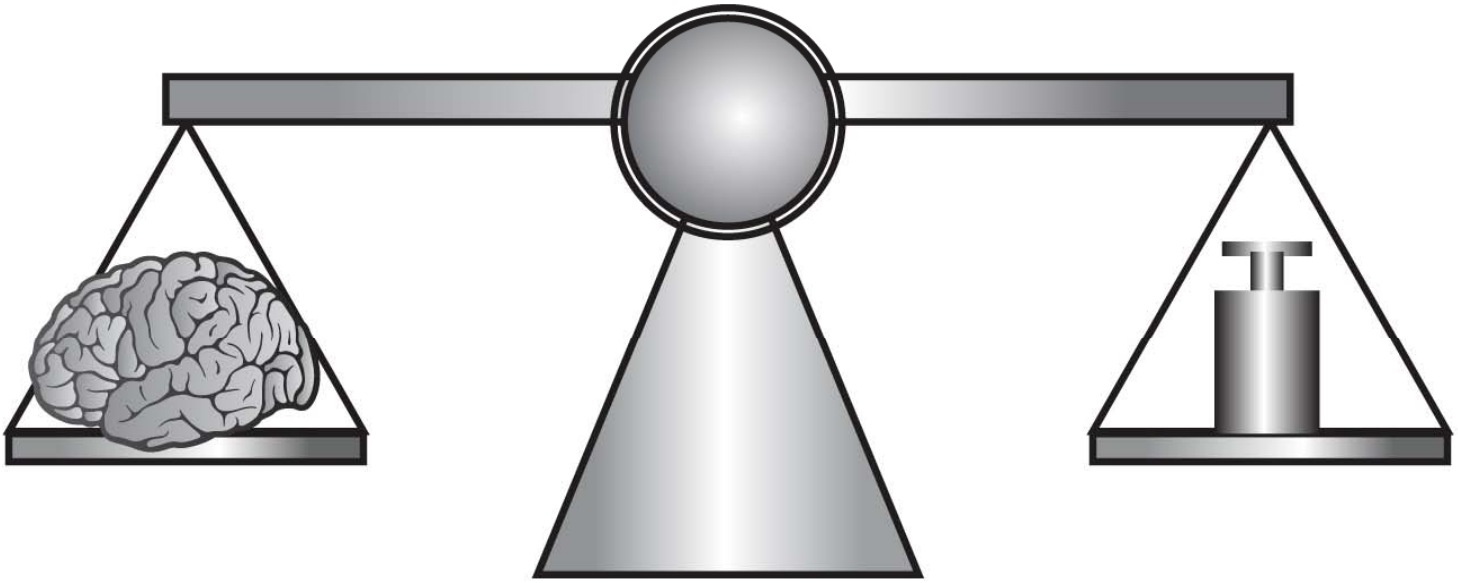
10 years

PerMIS



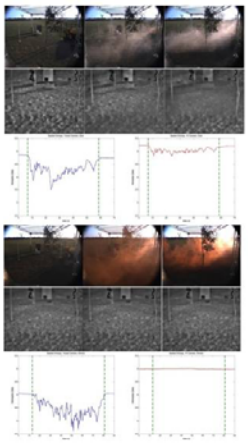
Digests

TUE-AM1
Measures & Metrics
Chairs: Hui-Min Huang & Seungbin Moon

Visual Metrics for the Evaluation of Sensor Data Quality in Outdoor Perception

Christopher Brunner and Thierry Peynot
Australian Centre for Field Robotics (ACFR), The University of Sydney

- This paper proposes an experimental study of quality metrics that can be applied to visual and infrared images acquired from cameras onboard an unmanned ground vehicle (UGV).
- The relevance of existing metrics in this context is discussed and a novel metric, called Spatial Entropy, is introduced.
- Selected metrics are evaluated on data collected by a UGV in clear and challenging environmental conditions, represented in this paper by the presence of airborne dust or smoke.



Proposals for New UGV, UMV, UAV, and HRI Standards for Rescue Robots

Dr. Robin R. Murphy
Center for Robot-Assisted Search and Rescue, Texas A&M University

- NSF-JST-NIST Workshop on Rescue Robots March 8-12, 2010 at Disaster City
  - 50 Participants from 16 universities in USA, Japan, and China
- Unmanned Ground Vehicles: *small diameter voids, vertical mobility*
- Unmanned Aerial Systems: *operating close to structures, response to failures and external disturbances, non line-of-sight operations, take-off and landing*
- Unmanned marine vehicles: *turbidity, station-keeping*
- Human-robot interaction: *perception, impact on human operator, robustness, interfaces*
- New classes of standards

Metric Selection for Evaluating Human Supervisory Control of Unmanned Vehicles (Uvs)

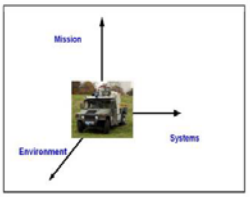
Birsen Donmez, University of Toronto
M. L. Cummings, MIT

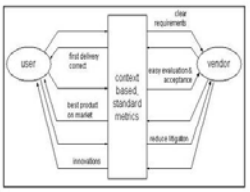
- The long term objective of this study is to develop principled methods that facilitate metric selection for evaluating human supervisory control of UVs.
- We previously identified a list of evaluation criteria that can help determine the quality of a metric, and generated a list of potential metric costs and benefits.
- In this paper, we conducted an experiment with subject matter experts.
  - We investigated which metric characteristics human factors practitioners consider to be important.
  - We also tested two different multi-criteria decision making methods (Analytic Hierarchy Process and Ranking Input Matrix) to help practitioners assign subjective weights to the cost/benefit criteria.
- The results show that
  - the majority of participants rated the evaluation criteria used in both tools as very useful.
  - the majority of participants' metric selections before using the methods were the same as the suggestions provided by the methods.

Performance Measures Framework for Unmanned Systems (PerMFUS): Models for Contextual Metrics

Hui-Min Huang, Elena Messina, Adam Jacoff--National Institute of Standards and Technoogy
Robert Wade-U.S.Army AMRDEC
Michael Mcnair-SAIC Company

- Contextual Metrics: Performance metrics should be associated with context to be meaningful.
- A set of generic metrics is established and is instantiated to the identified performance areas.
- Each performance area is identified with a set of specific metrics, each is associated with the context of autonomy and of environmental characteristics.
- A performance area--perfoamance of a subsystem--is affected by other subsystems. We established the cross-effects.
- Multiple adaptive lifecyels for the relatively new industry, UMS, are identified.

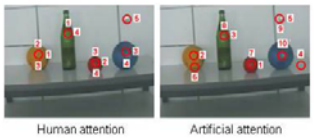


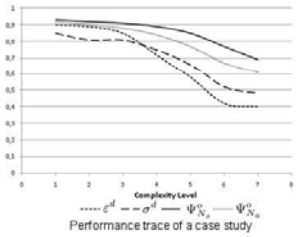


Towards Standardization of Metrics for Evaluation of Artificial Visual Attention

M. Zaheer Aziz, Bärbel Mertsching
GET LAB, Paderborn University, Paderborn, Germany

- Research in visual attention is important to achieve biologically plausible machine vision.
- Metrics and measurement methods to evaluate computational visual attention by comparing with human benchmark is discussed.
- Standardization is proposed at each involved step of evaluation:
  - Representation of focus of attention
  - Definition of equivalence between model and human focus of attention
  - Relation between attentional behaviors and evaluation metrics is established
  - Metrics for measuring efficiency, capability, and sequence proximity are proposed
- Trace using a dataset with increasing visual complexity can reflect achievement of models





Performance trace of a case study

Performance evaluation procedure for vision based object feature extraction algorithms

M. Kang, W. Choo, and S. Moon
Dept. of Computer Engineering, Sejong University, Korea

- This paper introduces performance evaluation procedure for vision based object feature extraction algorithms.
- The database, called OFEX, is constructed, reflecting the conditions a robot may encounter in the real environments: scale, illumination, pose variances, occlusion, and cluttered environments.
- The performance measures of recognition time, feature matching rate, recognition rate, and false acceptance rate are introduced.
- The experimental results, employing OFEX and performance measures, are discussed for a selected feature extraction algorithm.

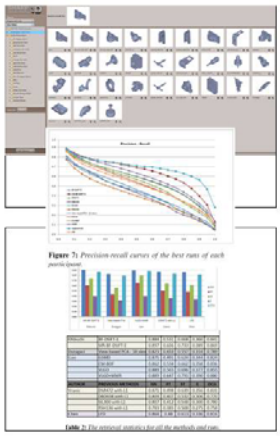
Environmental elements		conditions	
		Training set	Test set
Distance		1m	0.5, 1.5m
Illumination		200lx	60, 100, 400lx
Pose	Roll(x-axis)	0°	±20°
	Pitch(y-axis)	0°	±45°
	Yaw(z-axis)	0°	±45°
Occlusion		None	1/4, 2/4, 3/4
Clutter		No background	Complicated background

TUE-AM1 – Contd.
Measures & Metrics
Chairs: Hui-Min Huang & Seungbin Moon

Benchmarks, Performance Evaluation and Contests for 3D Shape Retrieval

Afzal Godil, Zhouhui Lian, Helin Dutagaci, Rui Fang, Vanamali T.P., Chun Pan Cheung
NIST

- Effectively searching a 3D repository has become an important area of research
- We will discuss the 3D shape retrieval benchmarks that we have developed and the different contests that we have organized
- Benchmarking allows researchers to evaluate the quality of results of different 3D shape retrieval approaches on a common dataset
- 3D shape Contests have a major impact on the 3D shape retrieval community and have advances the state of the art.
- Finally, we will discuss the different performance evaluation methods



TUE-AM2
Special Session I: Performance Metrics for Mixed Palletizing Operations
Organizers: Stephen Balakirsky & Henrik Christensen

Mixed Palletizing in Transformed Supply Chains: Operational Requirements and Technical Challenges

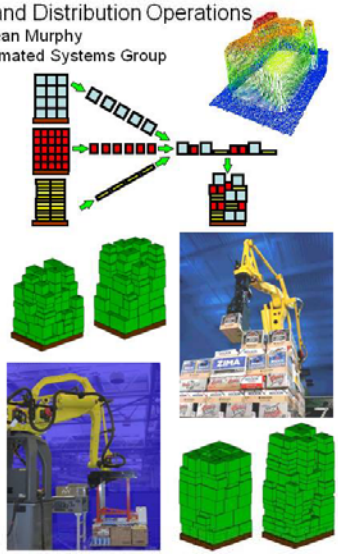
Larry M. Sweet, SVP Research & Development
C&S Wholesale Grocers

- Business perspective
  - Supply chain benefits
    - Warehouse and network wide
  - Customer benefits
  - Capital and operating costs
  - Operational constraints
- Technical challenges
  - Mixed case palletizing
    - Current capabilities
      - SKU diversity & complexity
      - Customer order characteristics
    - Future opportunities
  - System requirements

Industrial Robots in Warehousing and Distribution Operations

Don Faulkner and Sean Murphy
FANUC Robotics America, Automated Systems Group

- Field-proven applications for industrial robots in distribution are introduced
- Key concepts and practical considerations are reviewed
- Underlying technologies – robots, sensors, hardware, and software – are presented, including offline tools for system planning & design
- System features and limitations are compared and contrasted based on investment level
- Guidelines for identifying good candidate applications are identified



Planning in Logistics: A survey

Pushkar Kolhe, Henrik Christensen
Robotics and Intelligent Machines Center, Georgia Institute of Technology

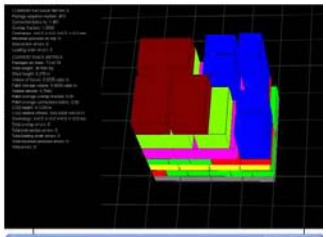
- The basis of logistics is a knapsack problem.
- We can theoretically determine the performance of a logistics system.
- Discuss algorithms to improve logistics.
- Bring together ideas from other research areas to improve planning in logistics.
- On-line mixed palletizing using dynamic and stochastic knapsack solutions.

Theoretical Benchmarks for Logistics

Metrics for Mixed Pallet Stacking

S. Balakirsky, T. Kramer, F. Proctor
National Institute of Standards and Technology

- Stacking objects onto pallets (mixed pallet stacking) accounts for over 60% of the volume of shipped good worldwide
- This paper addresses the question of "What makes a good pallet?"
- Pallet quality metrics are presented
- A tool for evaluating these metrics is also presented

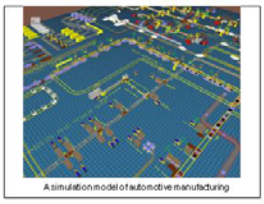


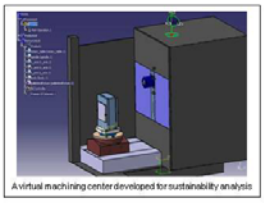
TUE-PM1
Interoperability & Sustainability
Chairs: John Horst & Ani Hsieh

Modeling and Simulation Analysis Types for Sustainable Manufacturing

D. Kibira, G. Shao, Y. T. Lee
Manufacturing Systems Integration Division, National Institute of Standards and Technology (NIST)

- The role of simulation in manufacturing is discussed
- The need to extend constructs within existing simulation applications to model sustainable manufacturing is discussed
- The paper classifies manufacturing simulation problems along: a) functional requirements and, b) data analysis
- The paper uses product lifecycle stage as a functional classification while data categories are discussed along the triple bottom-line
- An example of virtual machining requirements for sustainable manufacturing simulation is provided


A simulation model of automotive manufacturing


A virtual machining center developed for sustainability analysis

Metrics for the Cost of Proprietary Information Exchange Languages in Intelligent Systems

John Horst, Nathan Hartman, George Wong
NIST, Purdue University, The Boeing Company

The increasing number of intelligent software components is accompanied by an increasing number of proprietary information exchange languages between components. One of the challenges for the smart technology worker is to achieve intelligent system component interoperability, at the lowest cost possible, without sacrificing the freedom to choose from the entire spectrum of current and future software product offerings. We argue that this state of affairs is best achieved when correct, complete, and unambiguous information exchange standards are implemented in vendor products worldwide. If this is the common sense solution to information incompatibility costs and risks, why is standards-based interoperability so rarely seen?

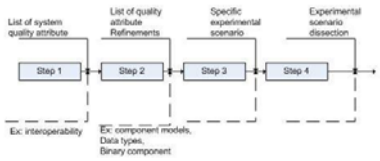
Corporate management is commonly not aware of the magnitude and seriousness of the interoperability problem and is therefore reluctant to commit to an investment in standards development. We argue that if there were a published set of realistic and comprehensive interoperability cost-risk metrics, which technology workers could employ to make an evidence-based business case, that would be sufficient to persuade management to support standards development.

Along the way, we point out that interoperability is directly and inextricably related to correct and complete communication between a sender and receiver, which implies that the proliferation of multiple, redundant (usually proprietary) languages (consisting of both format and meaning) is the heart of the problem.

Component Models in Robotics Software

Azamat Shakhimardanov, Nico Hochgeschwender, Gerhard K. Kraetzschmar
Univeristy of Applied Sciences Bonn-Rhein-Sieg, Germany

- Investigate component-oriented programming in robotics
- Discuss interoperability between robotics software frameworks
- Analyse four component-oriented robotics frameworks
- Assess them with respect to model-level interoperability
- Use four step approach for assessment
- Describe component modeling concepts used as evaluation criteria
- Identify possible levels and degrees of interoperability

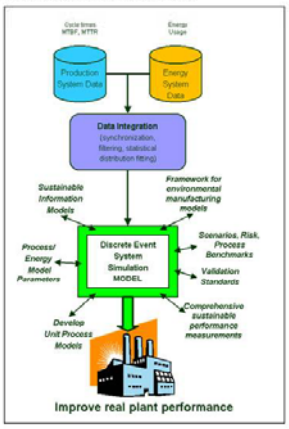


	Comp.	Data Flow Port	Service Port	Data Types	Conn.
OpenRTM	✓	✓	✓	✓	✓
Genom	✓	—	—	✓	—
ROS	✓	✓	—	✓	—
Orocos	✓	✓	✓	✓	✓

Benchmarking Production System, Process Energy, and Facility Energy Performance Using a Systems Approach

J. Arinez and S. Biller, General Motors
K. Lyons, S. Leong, G. Shao, B.E. Lee, and John Michaloski, National Institute of Standards and Technology

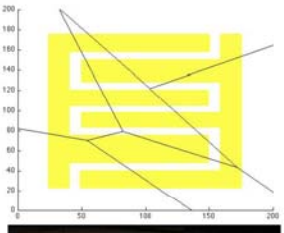
- Today, manufacturing energy and process data are loosely coupled.
- GM/NIST joint project to develop smarter process/energy integration
  - Develop standard methodology for developing integrated process/energy solutions
  - Determine energy savings/losses/risk avoidance/potential crisis points performance benchmarks
- Case Study
  - GM Precision Sand Casting production facility
  - Discrete Event Simulation (DES) model of integrated process/energy

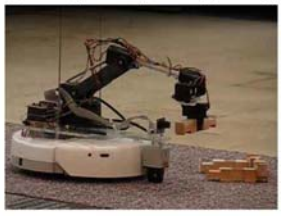


Complexity Measures for Distributed Assembly Tasks

M. Ani Hsieh and Joshua Rogoff
Drexel University, Philadelphia, PA 19104

- Existing work in distributed robotic assembly tasks focus mostly on the design of provable correct assembly rules
- This work considers the partitioning of the target structure to maximize assembly parallelization
- Mass-based partitioning strategies ensures equal allocation of assembly parts to robots but fail to take into the geometric complexity of the structure to be assembled.
- This work presents complexity measures to describe the complexity of the desired structure
- These measures will be used to improve existing allocation strategies

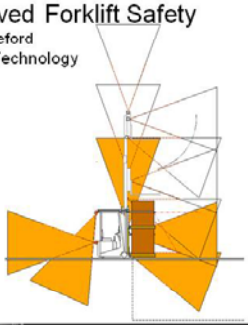


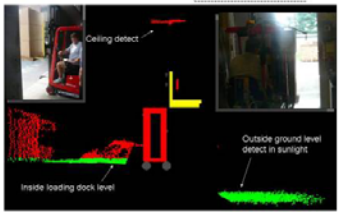


Advanced Sensing Towards Improved Forklift Safety

Roger Bostelman, Will Shackleford
National Institute of Standards and Technology

- OSHA estimates that there are 110,000 accidents each year with 80% involving a pedestrian
- $135 M immediate costs are incurred due to forklift accidents
- ANSI/ITSDF B56.5 Safety Standard [ANSI 2010] states that "a sensing device or combination of devices shall be supplied to prevent contact of the vehicle structure and installed equipment with people or objects appearing in the path of the vehicle in the main direction of travel."
- Advanced sensors are being studied to detect positive and negative obstacles and programmed to interoperate with/alert the forklift user to make safety decisions.





TUE-PM2
Special Session II: Unmanned and Autonomous Systems Test Technology
Organizers: Robert Heilman

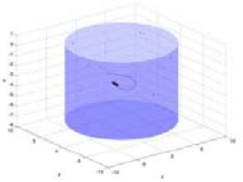
Unmanned and Autonomous Systems Test Technology

Rob Heilman

A Multi-Vehicle Tesbed for Underwater Motion Coordination

N. Sydney, S. Napora, D. Paley
Department of Aerospace Engineering, University of Maryland

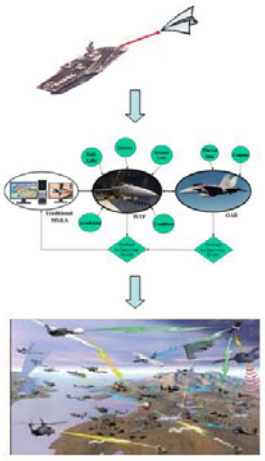
- The development of the Synthetic Collective Unmanned Underwater Laboratory (SCUUL) is presented
- The design a fleet of unmanned underwater vehicles (UUVs) is discussed
  - Six RC submarines with onboard sensors for inner loop control
- The test facility, the Neutral Buoyancy Research Facility, located at the University of Maryland, with its state-of-the art motion capture system is described
  - 367,000 gallon tank with 25 ft depth and 50 ft diameter
  - Motion capture system can output data in real time, or be saved for later analysis
- The research objective of the testbed is to apply dynamical systems theory and reduced-order modeling to cooperative UUV navigation, sampling, performance, and control
- The utility of the testbed to test UUVs and cooperative control algorithms is examined
  - Use of real time data from motion capture system in conjunction with onboard sensors in a feedback loop to evaluate algorithms



Measurement of Autonomous Operations

W. Hamel
USN, NAVAIR, Integrated Battlespace Simulation and Test (IBST)

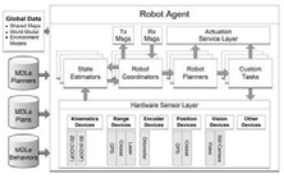
- Autonomous air systems moving from paper designs to flight
  - Basic Research
  - Applied Research
  - Future
- Building Trust in Autonomous Systems
- Autonomous Operations Test Methodologies
  - Constructive
  - Virtual
  - Live
- Risk Based Test Approach
  - V&V
  - Integrated Test
  - Based on Mission?
- Performance Metrics: Is accomplishing the mission sufficient?

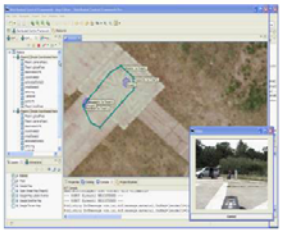


DCF® – A JAUS and TENA Compliant Agent-based Framework for UAS Performance Evaluation

N. Lenzi (nlenzi@i-a-i.com), B. Bachrach (bach@i-a-i.com), V. Manikonda (vikram@i-a-i.com)
Intelligent Automation Incorporated, 15400 Calhoun Dr. Suite 400, Rockville 20855

- The development of a JAUS and TENA compliant Distributed Control Framework (DCF) for controlling teams of dissimilar UASs is discussed.
- The functionality of graphical user interface (Vignette Editor) is discussed:
  - Visualizing the state of the UAS team
  - Creating T&E scenarios
  - Monitoring the UAS team performance
  - Generating automated T&E reports
  - Allows the user to issue real-time commands and to upload a new distributed control algorithm (mission) on-the-fly.
- The results of an evaluation of DCF by ARDEC personnel at Picatinny Arsenal in a multi-UGV, coordinated perimeter surveillance mission are discussed.

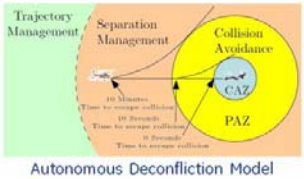




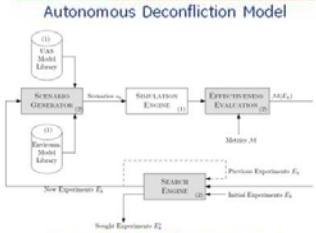
Testing and Evaluation Aspects of Integration of Unmanned Air Systems into the National Air Space

Mauricio Castillo-Effen and Nikita A. Visnevski
General Electric Global Research

- Technical Problem: Integration of UAS into NAS
  - Benefits: Economic, scientific, homeland security, safety, etc.
  - Hard problem: lack of data, no certifiable trust in autonomy, legacy systems, NAS transitioning to NextGen.
- Testing and Evaluation (T&E)
  - One of the tools, which may provide "level of trust" in the autonomous systems domain.
  - Provides highly needed data.
- Proposed T&E Planning Methodology
  - Design of Experiments Approach.
  - Uses Modeling and Simulation (M&S) to incorporate knowledge about the system.
  - Exploration of search space of input factors.
  - Metrics based on Measures of Effectiveness at the mission level.
  - Case studies:
    - Logistics.
    - Border patrol.



Autonomous Deconfliction Model



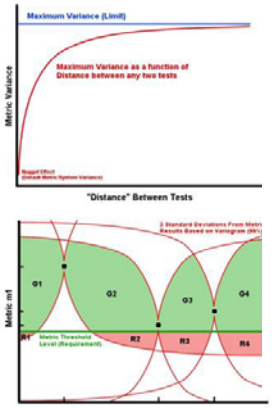
Automated Test Planning System

WED-AM1
Testing and Evaluating of Intelligent Systems
Chairs: Brian Weiss & James Gunderson

Evaluating Intelligent Systems with Performance Uncertainty in Large Test Spaces

Miles Thompson
Georgia Tech Research Institute

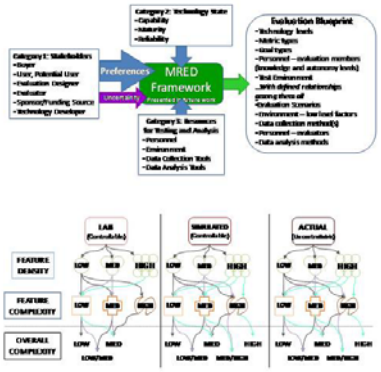
This presentation proposes a new methodology, the Spatial Variance Bounding Methodology, for designing tests/experiments for intelligent systems based on a meta-model method for characterizing a discontinuous, stochastic system. The methodology expands on the principles of Blind Kriging and spatial variance. The intent of these designs is to provide a framework for evaluating the performance of these intelligent systems across large test spaces that may take months or years to complete via traditional test/experiment design methods.



The Multi-Relationship Evaluation Design Framework: Creating Evaluation Blueprints to Assess Advanced and Intelligent Technologies

Brian A. Weiss[1] and Linda C. Schmidt[2]
National Institute of Standards and Technology[1], University of Maryland – College Park[2]

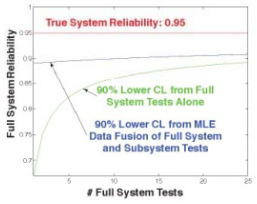
- MRED (Multi-Relationship Evaluation Design) is being designed as an automatic test plan generator
  - Output extensive evaluation blueprints customized to the technologies under test
  - Produce different blueprints over time as the technologies mature and inputs change
  - Acknowledge and factor in relationships and constraints among inputs and outputs
  - Address uncertain input from multiple evaluation stakeholders
- MRED is currently being verified with several pre-existing evaluation designs for validation

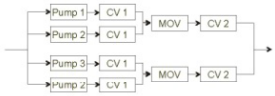


Implementation and Application of Maximum Likelihood Reliability Estimation from Subsystem and Full System Tests

Coire J. Maranzano and James C. Spall
Johns Hopkins University Applied Physics Laboratory

This paper provides an overview and examples of a novel and practical method for estimating the reliability of a complex system, with confidence regions, based on a combination of full system and subsystem tests. Maximum likelihood estimation is used to estimate the overall system reliability based on the fusion of system and subsystem test data. The method is illustrated on two real-world systems: an aircraft-missile system and a highly reliable low-pressure coolant injection system in a commercial nuclear-power reactor. The examples are used to demonstrate the some of the properties of the method.



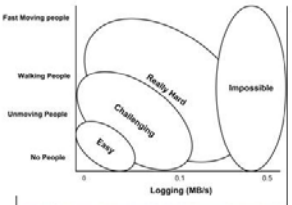


Example Non-Series System

"What do you do with a drunken robot?"
In Situ Performance Measurements of Intelligent Mobile Robots

J. P. Gunderson, L. F. Gunderson
Gamma Two, Inc.

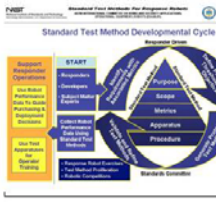
- Measuring the performance of intelligent robots requires careful design of the tests.
- We explore the impact of feedback loops on traditional testing approaches
- We use tests to probe the areas of stable performance, and measure performance degradation outside stable zones
- Movies of robots in action!



Comprehensive Standard Test Suites for the Performance Evaluation of Mobile Robots

Adam Jacoff, Hui-Min Huang, Elena Messina, Ann Virts, Anthony Downs
Intelligent Systems Division, National Institute of Standards and Technology

- Standard performance tests are being developed for response robots.
- The standards development cycle and philosophy is presented.
  - Driven by user-defined performance requirements, test methods are developed to measure robot capabilities in various categories.
  - Tests are iteratively refined, with user and developer feedback.
- A set of test methods is collected into a suite that collectively can be used to characterize a robot's performance and suitability for a particular application.
- The test suites are in various stages of prototyping, validating, and balloting.



Towards a Standardized Test for Intelligent Wheelchairs

Joelle Pineau[1], Robert West[1], Amin Atrash[1], Julien Villemure[1], François Routhier[2]
[1]Center for Intelligent Machines, McGill University
[2]Institut de réadaptation en déficience physique de Québec

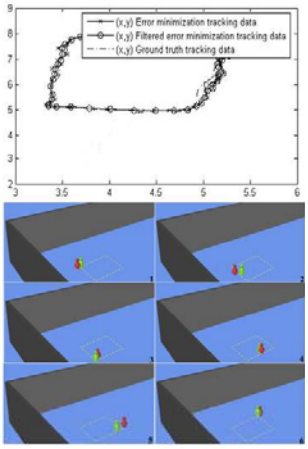
- We outline the development and validation of the SmartWheeler project, in particular the human-robot interface.
- We argue for the use of standardized testing in evaluating intelligent wheelchairs.
- We describe the Wheelchair Skills Test (WST), and discuss how to apply it to benchmark the performance and safety of intelligent wheelchairs.
- We present preliminary results from a user study which adapts the WST to evaluate the SmartWheeler's human-robot interface.

WED-AM1- Contd.
Testing and Evaluating of Intelligent Systems
Chairs: Brian Weiss & James Gunderson

Evaluation of Ultra-Wideband Technology for use in 3D Locating Systems

A. Kopp, K. Saidi, H. Khoury
Building and Fire Research Laboratory, National Institute of Standards and Technology (NIST), American University of Beirut

- A high powered (5 Watt peak power) ultra-wideband (UWB) based ranging system was used as the basis for 3D locating system.
- Localization algorithms were developed, which utilized trilateration, error minimization, multilateration, and low-pass filtering.
- Results obtained from algorithms were compared to a ground truth 3D locating system.
- Data obtained from UWB and ground truth systems were compared visually.
- Potential for UWB-based 3D locating systems for first responders and construction workers is discussed.
- Future research is suggested.



WED-AM2

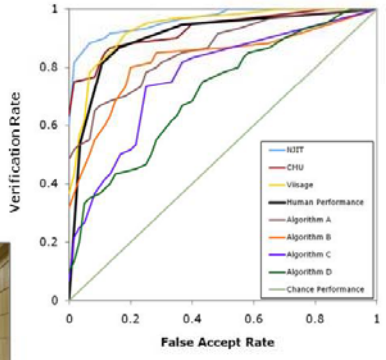
Special Session III: Evaluation of Human Detection and Tracking for Robot Safety, Collaboration and Interaction

Organizer: Tsai Hong & Roger Eastman

An Overview of Human and Machine Performance on Face Recognition

P. Jonathon Phillips
National Institute of Standards and Technology

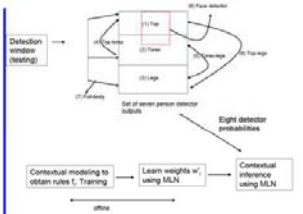
- Humans and algorithms matched face images taken under controlled and uncontrolled illumination.
- Machines can compete quantitatively humans.
- Recognition skills for humans and machines remain stable across changes in image quality
- Fusion of humans and machines can lead to higher recognition performance levels.



Detecting Humans under Partial Occlusion using Markov Logic Networks

Raghuraman Gopalan, and William Schwartz
University of Maryland, College Park

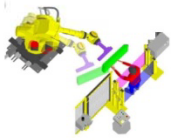
- Problem: Modeling the *semantic context* portrayed by humans with the surrounding scene to perform detection under partial occlusion.
- Context sources: Learning logical rules between detector probabilities of different human parts, and their relation with adjacent spatial windows.
- Performing probabilistic contextual inference in the framework of *Markov logic networks*.
- Learning the weights of the network parameters, under different occluding cases.
- Evaluations on standard datasets, and on dataset with occluding humans.



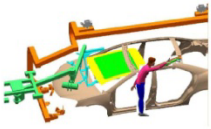
Transitional or Partnership Human and Robot Collaboration for Automotive Assembly

J. Shi, R. Menassa
Manufacturing Systems Research Lab, GM Global R&D Center, Warren, MI

- A near term and a practical strategy for robots to *complement,* instead of *emulate,* humans' capabilities was presented in order to achieve a shared goal in a shared context of manufacturing task execution.
- Two types, transitional and partnership, of human and robot collaborations for automotive assembly are defined, described, and examined.
- Detailed transitional and partnership human robot collaboration and their interaction mechanisms are studied to enable effective human and robot fail-safe interaction for a given action transition between a human and a robot.
- Critical new robot capabilities are identified that will ensure robots to "intelligently adapt" to the humans' presence and to establish the trust in a partnership collaboration.



Human and Robot Transitional Collaboration in Serially Executed Subtasks



Human and Robot Partnership Collaboration in Parallel Executed Subtasks

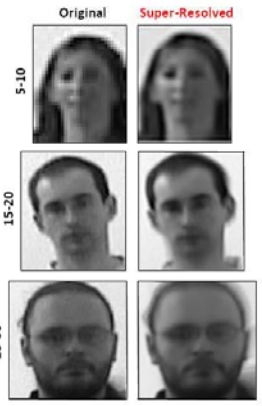
Performance Assessment of Face Recognition Using Super-Resolution

Shuowen Hu[1], Robert Maschal[1], S. Susan Young[1], Tsai Hong Hong[2], Jonathon P. Phillips[2]
[1]U.S. Army Research Laboratory, [2]NIST

Introduction: Performance of face recognition algorithms is dependent on a host of factors, one of which is resolution/scale. In surveillance applications, the number of face pixels acquired is typically very limited. Super-resolution reconstruction can form a higher resolution face image using a sequence of low-resolution frames, potentially aiding face recognition systems for improved performance.

Experiment: Parallel gait videos from the Database of Moving Faces and People (University of Texas at Dallas) is used for this work, in which subjects walk towards the camera starting from approximately 13 m away. Frames from three eye-to-eye distances (5-10, 15-20, and 25-30 pixels) corresponding to different distances from the camera are used to form super-resolved query sets. Original and super-resolved examples of subject faces are shown in the figure at the three different eye-to-eye distances.

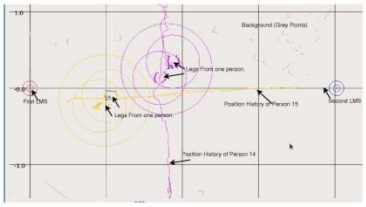
Face Recognition: The local region PCA baseline algorithm (Colorado State University) is used to assess performance using the original and super-resolved query sets. Performance assessment will be conducted using receiver operating characteristics (ROC) curves for each of the original and super-resolved query sets.



Inexpensive Ground Truth and Performance Evaluation for Human Tracking using multiple Laser Measurement Sensors

Will Shackleford,Tsai Hong, Tommy Chang
Intelligent Systems Division, National Institute of Standards and Technology (NIST)

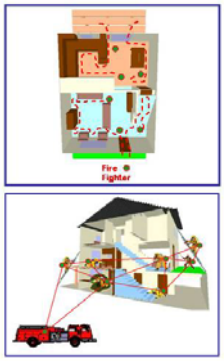
- INTRODUCTION
- Laser Measurement Sensor (LMS) MOUNTING
- DATA COLLECTION
- MANUAL VISUAL CALIBRATION
- BACKGROUND SUBTRACTION
- LEG AND HUMAN GROUPING
- TEST SETUP
- TEST RESULTS



Development of Performance Metrics and Test Methods for First Responder Location and Tracking Systems

Francine Amon[a], Camillo Gentille[b]
[a]Building and Fire Research Laboratory, [b]Information Technology Laboratory,
National Institute of Standards and Technology,

- Need: to develop performance requirements, metrics and test methods for location and tracking systems for first responders.
- A working group has formed to:
  - Collect and share information about their ongoing efforts,
  - Develop requirements for range of first responders
  - Characterize operational environments (physical, electromagnetic).
  - Determine scope of commonalities/differences in requirements and metrics,
  - Develop test methods that accommodate the needs of all/most first responders.
- The final deliverable: a draft document to be presented to appropriate standards developing organization(s).



This paper is authored by employees of the United States Government and is in the public domain. PerMIS'10, September 28-30, 2010, Gaithersburg, MD, USA. ACM 978-1-4503-0290-6-9/28/10

WED-AM2- Contd.

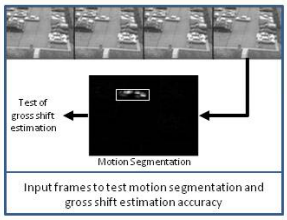
Special Session III: Evaluation of Human Detection and Tracking for Robot Safety, Collaboration and Interaction

Organizer: Tsai Hong & Roger Eastman

Automated Gross and Sub-pixel Registration Accuracy of Visible and Thermal Imagery

S. Won, S. S. Young, G. Seetharaman
Image Processing Branch, Army Research Laboratory

- The importance of accurate, sub-pixel image registration is discussed.
- The image segmentation and motion estimation methods are discussed.
- The use of visible and thermal imagery in the following tests:
  - Motion segmentation
  - Gross shift estimation with/out segmentation
  - Sub-pixel shift estimation of undersampled images
- Evaluation of registration through both segmentation and motion estimation is discussed as well as their performance



Input frames to test motion segmentation and gross shift estimation accuracy

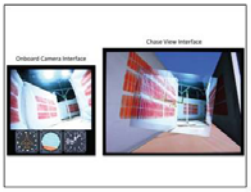
Undersampled images to test sub-pixel shift estimation

THU-AM1
Human-Machine Interaction and Collaboration
Chairs: Craig Schlenoff & Paul Oh

An Indoor Study to Evaluate A Mixed-Reality Interface For Unmanned Aerial Vehicle Operations in Near Earth Environments

James T. Hing, Justin Menda, Kurtulus Izzetoglu, Paul Y. Oh
Drexel University, Philadelphia, PA

- Discussed are efforts to increase pilot situational awareness (SA) using a mixed reality chase view piloting interface for direct control of UAVs.
- A series of UAV piloting experiments were performed using a flight simulation package, UAV hardware, and an indoor, six degree of freedom, robotic gantry.
- Performance data was collected and cognitive workload during tests was captured using the NASA task load index and a near infrared Spectroscopy (fNIR) system.
- Behavioral analysis showed that the chase view interface improved pilot performance in near Earth flights and increased their SA.
- fNIR analysis showed that a subjects cognitive workload was significantly less while using the chase view interface.

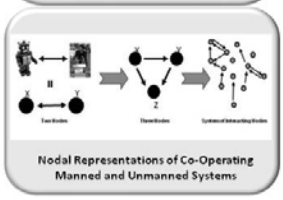


A Network-based Approach for Assessing Co-Operating Manned and Unmanned Systems (MUMS)

Dr. Lora G. Weiss
Georgia Tech Research Institute

- Traditionally, robots have been programmed to do precisely what their human operators instruct them to do.
- More recently, they have become more sophisticated, intelligent, and autonomous.
- Once they reach a sufficiently high level of intelligent autonomy, they can support more collaborative interactions with each other and with people.
- As robots become more and more intelligent, we will begin designing systems where robots interact with humans, rather than designing robots that are commanded by people with continual oversight.
- One approach to assessing how humans and robots will interact in the future is to frame the problem as a collection of intelligent nodes.
- Multiple, collaborating, and interacting manned and robotic systems can be represented as a collection of dynamic, interacting nodes.
- This paper develops preliminary metrics to support understanding the extent of preferential attachment that would arise in a system of co-operating manned and unmanned systems (MUMS).
- The metrics seek to help explain if attachments are localized to specific situations or if they are more pervasive throughout a MUMS society.


Graphical Concept of Many Co-Operating Manned and Unmanned Systems


Nodal Representations of Co-Operating Manned and Unmanned Systems

Lessons Learned In Evaluating DARPA Advanced Military Technologies

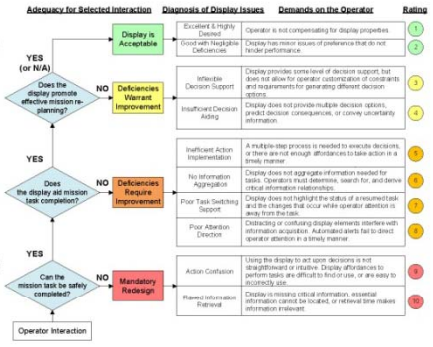
Craig Schlenoff, Brian Weiss, Michelle Steves
National Institute of Standards and Technology

- For the past six years, NIST have served as the Independent Evaluation Team for two major DARPA programs.
- DARPA ASSIST's goal is to exploit soldier-worn sensors to augment a Soldier's situational awareness, mission recall and reporting capability
- DARPA TRANSTAC's goal is to demonstrate capabilities to rapidly develop and field free-form, two-way speech-to-speech translation
- Both of these efforts are concluding and as such this paper will focus on overall lessons learned in evaluating these types of technologies.

Modified Cooper Harper Scales for Assessing Unmanned Vehicle Displays (MCH-UVD)

Birsen Donmez, University of Toronto
M. L. Cummings, MIT
Amy S. Brzezinski, NASA Johnson Space Cen.
Hudson D. Graham, MIT

- We propose a subjective UV display evaluation tool: MCH-UVD
- MCH-UVD is adapted from the Cooper-Harper aircraft handling scale by shifting focus to support of operator information processing.
- An experiment was conducted to evaluate and refine MCH-UVD, as well as assess the need for mission-specific versus general versions.
- Participants (86%) thought that MCH-UVD helped them identify display deficiencies, and 32% said that they could not have identified the deficiencies without the tool.
- No major additional benefits were observed with mission-specific versions over the general scale.

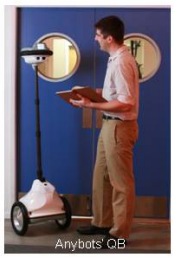


Using the "Negative Attitude Toward Robots Scale" with Telepresence Robots

Katherine Tsui[1], Munjal Desai[1], Holly Yanco[1], Henriette Cramer[2], and Nicander Kemp[3]
[1]University of Massachusetts Lowell, [2]SCIS & Mobile Life Centre, [3]KZA b.v.

- What is NARS?
  - Developed in 2003 by Nomura et al. in Japan
  - Performance measure for people's attitudes towards robots (interaction, social, emotion)
- Consistency validation of English NARS
- Is NARS appropriate for use with telepresence robots?
  - Used QB and Vgo robots
  - Online video study
  - People operating telepresence robots

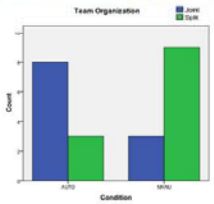

Vgo Communications' Vgo


Anybots' QB

Teams Organization and Performance Analysis in Autonomous Human-Robot Teams

Huadong Wang, Michael Lewis, Shih-Yi Chien
School of Information Sciences University of Pittsburgh

- ❑ This study addresses the interaction between automation and organization of human teams in controlling large robot teams performing an Urban Search and Rescue (USAR) task.
- ❑ "Shared Pool" conditions which 24 simulated robots were controlled by teams of 2 participants
- ❑ Autonomy vs. Manual Control in robot navigation
- ❑ Three self-organizing team strategies in this shared pool condition were identified: joint control, mixed control, and split control.
- ❑ Automating path planning improved system performance such as victims found and explored area.
- ❑ Manual condition participants were less likely to choose joint control than were Autonomy participants.
- ❑ Effects of team organization favored operator teams who shared authority for the pool of robots.



THU-AM1 – Contd.
Human-Machine Interaction and Collaboration
Chairs: Craig Schlenoff & Paul Oh

Intentions and Intention Recognition in Intelligent Agents

Dr. Gary Berg-Cross — Knowledge Strategies — gbergcross@gmail.com
Dr. Christopher Crick — Brown University — chriscrick@cs.brown.edu

Intentions and intention recognition (IR) is an important component of Theory of mind (ToM) & the ability to infer others' internal intention based on their behaviors.

- In the social-perceptual sphere intention recognition allows the same agent movements to be understood as "moving", "giving" or "loaning" an object.

Both primate cognition and child development studies provide validation of its role and ontogenesis in integrating actions & perception within dynamic goals.

We outline studies of the neuro and psych-physiology investigation of intentions and IR.

Following this we discuss some efforts to apply these ideas to cognitive robots and sketch out one developmental architecture that indicates how such a capability may emerge.

Over time we have seen a movement from deliberative robots, rooted in early AI model, to one of reactive robots, followed by behavior-based robots and more recently intentional and motivational robots but we are still early in the work of development of computational models of intention-recognition.

Nevertheless the convergence of work from animal, child development and cognitive robotics is encouraging.

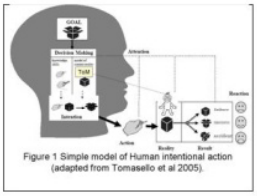
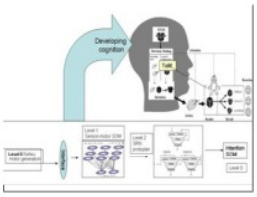


Figure 1 Simple model of Human intentional action (adapted from Tomasello et al 2005).

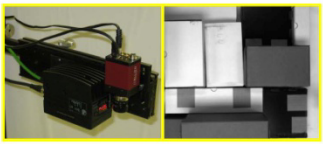


THU-AM2
Evaluation of Sensors for Object Pose Estimation in Manufacturing Applications
Organizers: Roger Eastman, Tsai Hong & Hui-Min Huang

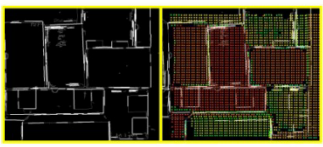
6DOF Pose Estimation
The Need for Standardization in Industrial Applications

Shaun Edwards, Clay Flannigan, Paul Evans
Southwest Research Institute

- A Position Paper is Presented
- The Evolution of Industrial Sensing from 2D (3DOF) to 3D (6DOF) is Discussed
- Examples of Industrial-Focused Sensors are Given
- Challenges with Non-Standard Sensor/System Performance Metrics are Identified
- Current Work on a Representative Application is Presented
- Potential Approaches to Standardized Testing are Discussed



Example Industrial Application: 2D and 3D Imager Fused to Sense a Cluttered Environment (Random Boxes)

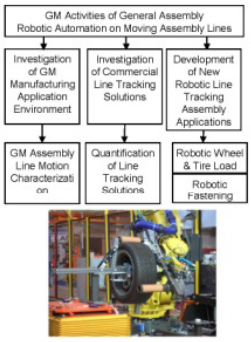


Classical 2D Edge Finding, Enhanced with 3D Range Information (Colorized Height)

Flexible Robotic Assembly in Dynamic Environments

J. Shi, R. Menassa
Manufacturing Systems Research Lab, GM Global R&D Center, Warren, MI

- GM's comprehensive effort in achieving robotic flexible assembly on moving assembly lines is first presented.
- Three types of assembly methods in automotive general assembly are summarized, and assembly alignment and manipulation motion characteristics in each type of the assembly methods are examined.
- In order to achieve robust and flexible robotic assembly in an unfixtured and dynamic manufacturing environment, several key relationships are studied and presented:
  - Required robotic manipulation in relationship with assembly methods,
  - Required robotic perception skills in relationship with assembly part motion,
  - Required robotic assembly system robustness in relationship to assembly tolerance and dynamic uncertainty.



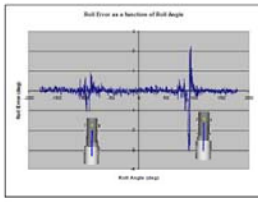
GM Activities of General Assembly Robotic Automation on Moving Assembly Lines

Smart Sensing for Real-Time Pose Estimation, Assembly, and Inspection Using 3D Laser Scanning Systems

Chad English
Department of Research and Development, Neptec Design Group

- State-of-the-art Pose Estimation systems are discussed, especially for AR&D
- Autosynchronous scanning sensors are compared to flash and MEMs lidars
- Smart scanning is defined using the "more information, less data" (MILD) paradigm to optimize efficient pose estimation
- Pose estimation is applied to various dynamic operations, including assembly, metrology, object recognition, inspection & digitization, mapping, navigation, and GIS.
- Benefits of closing the kinematic loop using pose estimation are discussed and applied.
- Challenges for performance metrics and testing are discussed with examples.

Real vs virtual shuttle views from pose estimate.



Dynamic Performance Evaluation of 6D Laser Tracker Sensor

Kam Lau, Yubing Yang, Yuanqun Liu, Henry Song
Automated Precision Inc.

• Introduction to API 6D laser tracker sensor dynamic performance evaluation. API product I360 will be used to demonstrate the evaluation method. The tracker i360 system provides information of r, α, β, pitch, yaw and roll. r, α, β are the polar coordinates of the i360 measured by Tracker, and pitch, yaw and roll are three orientation angles measured by i360 itself.

• The oscillating linear stage method is successfully employed for dynamic evaluation of the i360 system. Each of the 5 angle measurement of the 6D sensor system was evaluated with the oscillating linear stage method, schematically illustrated in Fig1.

• Test results with Oscillating linear stage method employed for API I360 sensor dynamic evaluation before and after dynamic compensation showed that the presented dynamic evaluation method is effective and accurate.

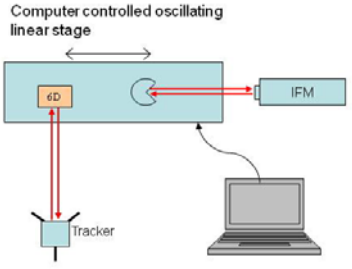


Fig.1 Oscillating linear stage method

Methodology for evaluating Static six-degree-of-freedom (6DoF) Perception Systems

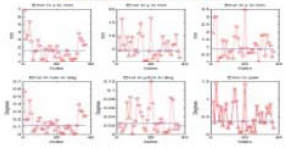
Tommy Chang, Tsai Hong, Michael Shneier, Mili Shah, Roger Eastman
National Institute of Standards and Technology and Loyloa University Maryland

Goal:
Toward evaluation of performance for static 6DoF perception systems.

Application of two fundamental approaches toward evaluating a static, vision based, six-degree-of-freedom (6DoF) pose determination system that measures the position and orientation of a part.

- The first approach uses groundtruth carefully obtained from a laser tracker
- The second approach avoids using any external groundtruth.

The evaluation procedure focuses on characterizing both the system's accuracy as well as its precision.



Std Dev	x (mm)	y (mm)	z (mm)	Rx (deg)	Ry (deg)	Rz (deg)
CondA	0.038	0.025	0.070	0.0344	0.0643	0.0096
CondB	0.046	0.018	0.078	0.0409	0.0521	0.0051
CondC	0.052	0.137	1.043	1.8300	0.3541	0.0622

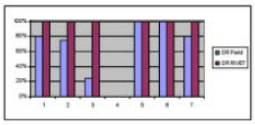
THU-PM1
Special Session V: Integrated Performance Assessment Through Experimentations
Chairs: Marshal Childers

Evaluating Hierarchical Planner Performance through Field Experiments and Simulation

Juan Pablo Gonzalez,
General Dynamics Robotic Systems

Marshal Childers, Barry A. Bodt
Army Research Laboratory

- This paper presents the results of using simulation combined with field testing in order to quantify the performance of integrated hierarchical planners
- The results presented here indicate that although there are few theoretical guarantees in integrated planners, significant trends do show up when evaluating them in both simulation and the field.
- By leveraging both simulation and field tests it is possible to obtain better performance metrics for integrated planners and unmanned vehicles in general.

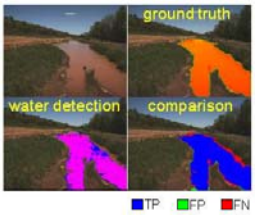


Percentage of completed runs

Evaluating the Performance of Unmanned Ground Vehicle Water Detection

Arturo Rankin, Tonislav Ivanov, Shane Brennan
Jet Propulsion Laboratory (JPL), California Institute of Technology

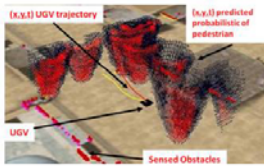
- UGV water detection algorithms were developed during the RCTA program.
- A method for evaluating the performance of water detection is discussed.
  - A software tool for ground truthing water in image space was developed using OpenGL and GIMP.
  - The user selects vertices in the first frame.
  - Edge detection is used to trace the water's edge.
  - The 2D location of vertices are automatically updated from frame to frame using stereo vision.
  - True positive (TP), false positive (FP), and false negative (FN) water detection rates are discussed.
- A method for evaluating the performance of localizing detected water is discussed.
  - Methods for ground truthing water in map space are discussed.
  - Several metrics for characterizing water localization performance are discussed.



Autonomous mobility for areas with large number of pedestrians

Alberto Lacaze, Karl Murphy, Nenad Uzunovic, Joseph Putney
Robotic Research LLC

- The need for autonomous SUGVs in crowded scenarios is discussed.
- An architecture for control is presented for a dynamic environment where the classical separation of perception and planning is not easily accomplished.
- The representation utilized for control is outlined.
- Prediction mechanism is presented to:
  - Provide accurate predictions in the 2-10 s area
  - Provide (x,y,t) planning support
  - Introduce the alpha-beta problem
- Current results and experimentation of architecture will be presented.



Observations on Single Operator Performance Controlling Two UGVs in a Field Assessment

Susan G. Hill
U.S. Army Research Laboratory, Human Research & Engineering Directorate

- An experimental field assessment that examined "span of control" is discussed.
- "Span of control" addresses the question of how many unmanned vehicles can be successfully operated simultaneously by an operator.
- Three operators ran two trials each, with each controlling two unmanned ground vehicles simultaneously.
- Observations that relate to the ability to assess operator span of control in a field test environment are discussed.

Assessing Unmanned Ground Vehicle Tactical Behaviors Performance

Marshal Childers, Barry A. Bodt
Army Research Laboratory

Richard Camden
General Dynamics Robotic Systems

- This paper provides an overview of our technology assessment process with a focus on performance metrics to foster collaboration with other investigators in order to further derive meaningful metrics for UGV autonomous navigation.
- For select experiments we present the enabling technologies, experimental factors and design, some exemplar response measures, and a summary of pertinent technology assessments.
- If measures are to be meaningful with respect to assessing capabilities in tactical behaviors, the metrics used to evaluate system performance must take into account end-to-end performance in a relevant environment

Tactical Behavior Challenges

Count of terrain challenges solved (Gold course)

Blocks	No blocks			Two blocks (1 block@exit & 1 road block)			Four blocks (2 blocks@exit & 2 road blocks + path to hole)		
Planner	ADH	DR	Mono	ADH	DR	Mono	ADH	DR	Mono
Road network only (prior) data	1/2	2/2		0/2	4/4			6/7	2/2
STM + road network (prior) data	0/1	1/2		0/1	1/2			8/10	5/5
Binary + road network (prior) data	2/2	2/2		3/4	4/4			10/10	23/24

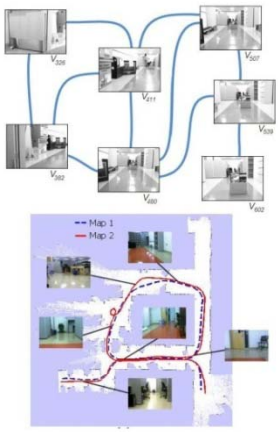
Terrain Challenges Solved

THU-PM2
Special Session VI: Performance Evaluation for Mapping and Navigation in Unstructured Environments
Organizers: Rolf Lakaemper & Raj Madhavan

Evaluation criteria for appearance based maps

Gorkem Erinc, Stefano Carpin
University of California, Merced, USA

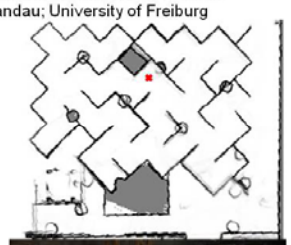
- Performance evaluation criteria to measure the quality of appearance based maps
- Task-based evaluation
  - Localization
  - Planning
  - Navigation
- Localization: accuracy vs coverage
- Planning: exactness vs completeness
- Validation of proposed metrics



Evaluation of Maps using Fixed Shapes: The Fiducial Map Metric

Sören Schwertfeger, Adam Jacoff, Chris Scrapper, Johannes Pellenz, Alexander Kleiner
Intelligent Systems Division, National Institute of Standards and Technology (NIST);
The MITRE Corporation; University of Koblenz-Landau; University of Freiburg

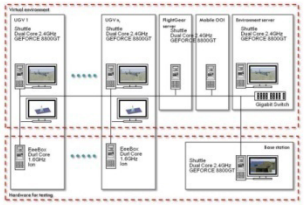
- Map metric measures different quality attributes like: Coverage; Global/ Relative Accuracy and Local Consistencies
- Artificial features (round shapes) placed in environment
- Experiments: two half-barrels on either side of a wall
- Error estimation using (relative) positions of fiducials in map vs. ground truth positions
- Experiments using maps generated at 2010 NIST Response Robot Evaluation Exercise at Disaster City
- Algorithm:
  - Low demand on ground truth data
  - Robust against most map errors
  - Humans can easily estimate map quality by checking the fiducials

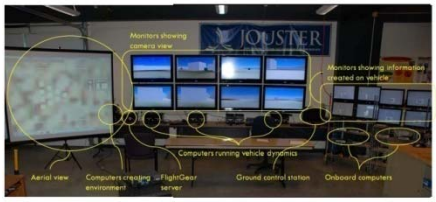


The Platform- and Hardware-in-the-loop Simulator for Multi-robot Cooperation

T. Furukawa, L.C. Mak, K. Ryu and X. Tong
Virginia Polytechnic Institute and State University

- Platform- and Hardware-in-the-loop Simulator (PHILS) for multi-robot cooperation;
- Consists of computers with GPUs, monitors, network switch and server-client simulation software;
- Allows tests and evaluations of cooperative performance and hardware performance;
- Implements synchronous and asynchronous communication and communication delay and loss.





World Modeling for Autonomous Navigation in Unstructured and Dynamic Environments

Rolf Lakaemper, Temple University
Raj Madhavan, University of Maryland, College Park & NIST

- How to create and experimentally validate a world modeling framework for unstructured environments amidst dynamic objects
- Focus is on development of truly automated guided vehicles (AGVs) for warehouses and factory floors
- Scientifically sound and statistically significant metrics, measurement, and evaluation methodologies for quantifying the AGV performance are required