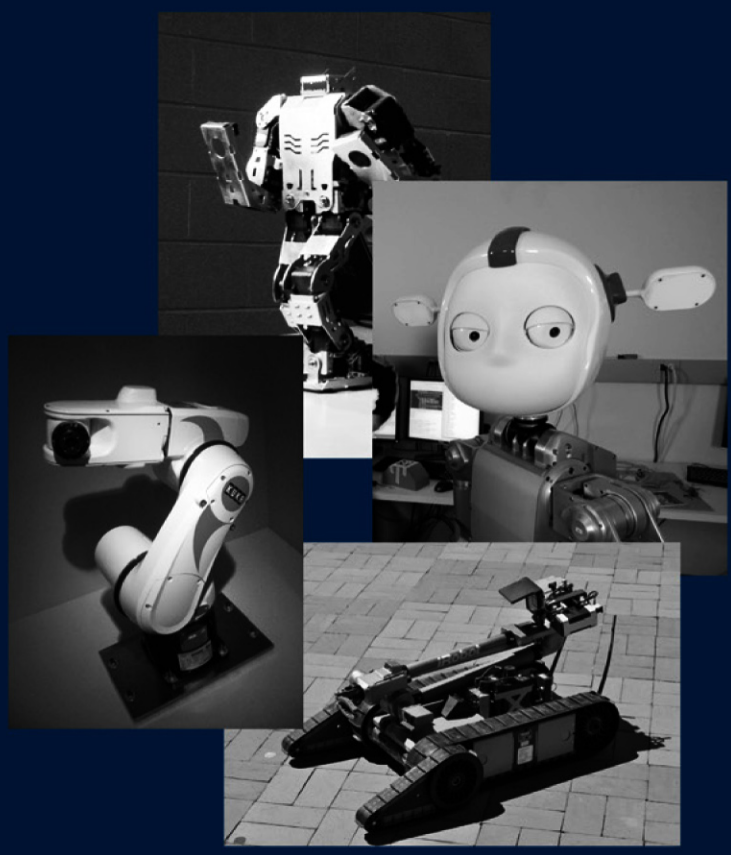
RESEARCH

ROBOTICS

RIM

KEEPS YOU ON THE EDGE



RIM @ GEORGIA TECH RESEARCH PORTFOLIO 2011

ENGINEERING BIOMEDICAL ELECTRICAL COMPUTER MECHANICAL AEROSPACE INTERACTIVE COMPUTING

INNOVATING THE FUTURE OF ROBOTICS

The Center for Robotics and Intelligent Machines (RIM@Georgia Tech) is helping define the new face of computing through a unique emphasis on education and research in robotics. The Center positions Georgia Tech to become a world leader in these promising, revolutionary new technologies.

RIM@Georgia Tech's activities leverage the strengths and resources of Georgia Tech by reaching cross traditional boundaries to embrace a multidisciplinary approach. The College of Computing, College of Engineering and the Georgia Tech Research Institute play key, complementary roles through Tech's traditional expertise in interactive and intelligent computing, control, and mechanical engineering.

Emphasizing personal and everyday robotics as well as the future of automation, faculty involved with RIM@Georgia Tech help students understand and define the future role of robotics in society.n addition, well—established industry relationships provide a path for technology transfer and commercialization - a crucial objective for RIM@Georgia Tech projects.

Basic and applied research is the heart of RIM@Georgia Tech. The study of basic engineering problems in robotics is central to our work, but equally important is the integration of innovation and discoveries into real-world systems and applications RIM@Georgia Tech research encompasses more than a dozen laboratories and 30-plus faculty from many different engineering disciplines.

Moreover, coordinating robotics research through the RIM@Georgia Tech center focuses those efforts and promotes development of a unified research strategy. This is an important step toward our goal of becoming one of the world's premiere robotics research institutions.

TABLE OF CONTENTS

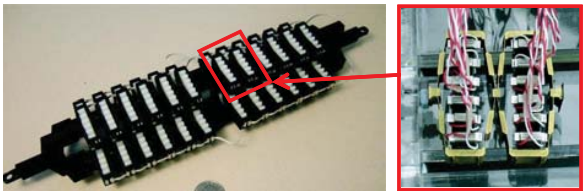
Fingerprint Method for Stochastic Arrays

Bio-Inspired Robotics and Human Modeling Laboratory
David MacNair, Ph.D. Candidate; Jun Ueda, Faculty Advisor
George W. Woodruff School of Mechanical Engineering

This research was supported by:
NSF Cyber-Physical Systems
ECCS-0932208.

Introduction

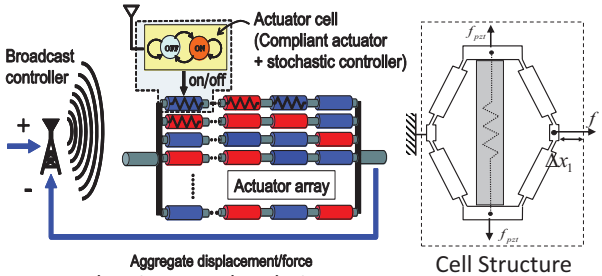
Generation of natural movements, or the movements created by biological systems, has been one of the biggest scientific questions discussed in physiology for decades. Biological muscle is known to have properties of stochasticity compliancy, and redundancy. Building these properties into artificial muscles, or actuator arrays, hopefully will allows robotic devices to move and manipulate objects in human environments effectively and in a way that humans observers feel comfortable around.



Stochastic Actuator Array | Individual Cells

Stochastic Cells

The actuator arrays are built from small actuator cells with structural elasticity. These cells are controlled using a bistable stochastic process wherein all cells are given a common input probability (control) value which they use to determine whether to actuate or relax. This greatly simplifies wiring and control, and greatly reduces the effect of actuator hysteresis. The structural elasticity allows the cells to be combined into complex arrays and for the resultant array to appropriately interact with the environment.



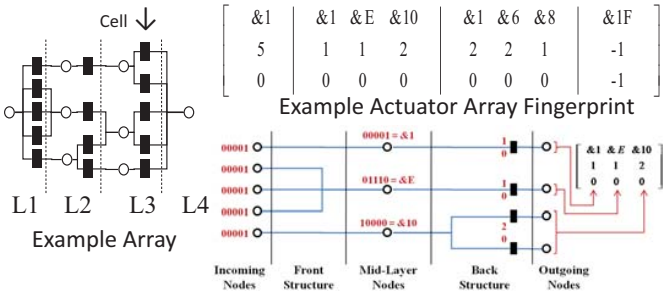
Stochastic Control Technique

Actuator Array Force

The linearity and constant slope for each actuator array's input probability (control) to mean force output allows a controller to directly command a desired mean force as long as the minimum and maximum forces at the endpoints of the array's travel have been identified and the array's current length is known. Additionally, when combining displacement and force sensors with the actuator array, the current belief of a manipulator and its payload can be continuously updated by higher level controllers and learning techniques. With these sensors, the array can also be continuously calibrated so that it remains robust and accurate despite multiple cell failures. This robustness characteristic may prove very useful in remote environments where a robot still needs to finish a task and move to a safe position despite multiple actuator failures.

Actuator Arrays

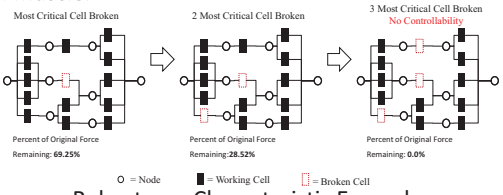
Cells can be arranged into complex topologies giving a wide range of strength, displacement, and robustness characteristics. Actuator arrays can be represented simply and compactly using a layer-based description, or fingerprint, which uses hexidecimal numbers to represent complex structure and decimal numbers for cell and rigid link connections.



Example Array Layer 2 Fingerprint Details

Fingerprint Method

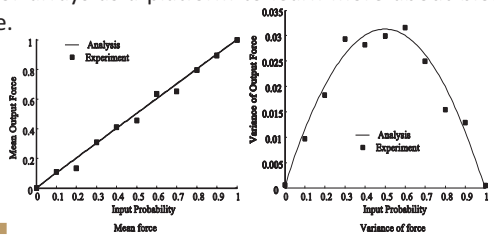
Infinite possible topologies exist, but given intelligent engineering design restrictions the fingerprint method can describe and automatically generate every possible stochastic actuator array topology for a given number of cells, and can calculate actuator array properties such as: travel, required actuator strength/displacement, force range, force variance, and robustness for any array topology. These properties are necessary when designing arrays for use in robotic applications and provide interesting similarities with biological muscle.



Robustness Characteristic Example
3 Cells Lost Gives No Controllability

Actuator Array Variance

The relationship between the input probability and variance is quadratic and does not monotonically increase, while the variance is expected to increase monotonically and proportionally with respect to command input in biological systems. Maximum voluntary force for a biological muscle, however, may not be the maximum possible for that muscle, but rather a result of limits on the control signal. This would be similar to limiting the actuator array input values, for example to p=0 ~ p=0.4. This is likely due to muscular fatigue and provides an interesting avenue for future research using actuator arrays as a platform to learn more about biological muscle.



Robotics@GT & Intelligent Machines

Pneumatically-Powered Robotic Exoskeleton to Exercise Specific Muscles in Space

Bio-Robotics and Human-Modeling Laboratory: Greg Henderson, M.S. Student, Jun Ueda, Faculty Advisor
George W. Woodruff School of Mechanical Engineering

SMART Scholarship
Science, Mathematics & Research for Transformation

Introduction

As humans inevitably extend their domains to the reaches of space; either traveling to Mars, etc., they must account for the effects of micro-gravity on their bones and muscles. In space, without the constant force of gravity, muscles atrophy and bone loss occurs. To mitigate these effects, a wearable robotic exoskeleton is proposed to allow a human to exercise while in a micro-gravity environment.

Conceptual futuristic space Shuttle for venturing further into depths of space[1]

Cosmonauts on stretchers due to muscle atrophy from staying aboard space station Mir[2]

Objectives / Motivation

- Exercise Anti-Gravity Muscles (Legs and Lower back) to shorten rehabilitation time when returning to a higher gravity environment
- Minimize Bone Losses since some loss may be permanent
- Preventing injuries, such as bone fractures, which could jeopardize missions such as to Mars where medical help is limited

Current Technology/Limits

Treadmill (with Elastic Bands over shoulders to simulate gravity)[3]

Current Technology:

Treadmill, Cycle, Lower Body Negative Pressure Device (LBNP), etc.

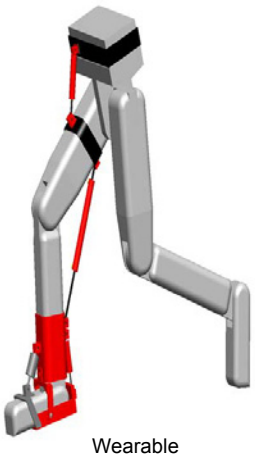
Limitations:

− Geometrically Large and Heavy - Physical space constraints and cost to send into space
− Time Commitment - Series exercise; cannot actively work on mission-related tasks while exercising

Wearable Exoskeleton

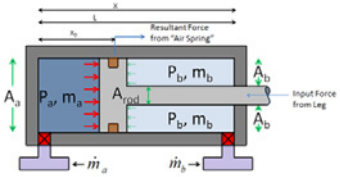
Advantages over Current Technologies:

+ Light-weight
+ Comparatively inexpensive
+ Parallel Exercise exercising while doing other activities)
+ Modular Design
+ Adjustable parameters such as spring constants, damping coefficients, etc.



Wearable Exoskeleton

Dynamic Modeling



Pneumatic Actuator and Valves

Model of Pneumatic Actuator

$$F_{(Hip,\ Knee,\ Ankle)} = P_a A_a - P_b A_b - P_{atm} A_{rod}$$

$$\dot{F}_{(H,K,A)} = \dot{P}_a A_a - \dot{P}_b A_b$$

$$\dot{F}_{(H,K,A)} = \frac{RTA_a}{V_a}\dot{m}_a - \frac{RTA_b}{V_b}\dot{m}_b - \frac{P_a A_a}{V_a}\dot{V}_a + \frac{P_b A_b}{V_b}\dot{V}_b$$

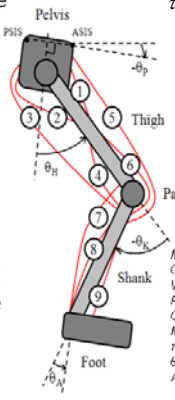
+

Combined Model (Leg and Pneumatic Actuator)

$$M\begin{bmatrix}\ddot{\theta}_H\\ \ddot{\theta}_K\\ \ddot{\theta}_A\end{bmatrix} = C\begin{bmatrix}\dot{\theta}_H^2\\ \dot{\theta}_K^2\\ \dot{\theta}_A^2\end{bmatrix} + V\begin{bmatrix}\dot{\theta}_H\dot{\theta}_K\\ \dot{\theta}_H\dot{\theta}_A\\ \dot{\theta}_K\dot{\theta}_A\end{bmatrix} + P\begin{bmatrix}\ddot{x}\\ \ddot{y}\end{bmatrix} + G + \begin{bmatrix}M_H\\ M_K\\ M_A\end{bmatrix} - \begin{bmatrix}\tau_H\\ \tau_K\\ \tau_A\end{bmatrix}$$

$$\tau_{MUSCLES} = \begin{bmatrix}M_H\\ M_K\\ M_A\end{bmatrix} = A(\theta)^{i\times n} f^{n\times 1}$$

$$\tau_{ACTUATOR} = \begin{bmatrix}\tau_H\\ \tau_K\\ \tau_A\end{bmatrix} = B(\theta)^{3\times 3} F^{3\times 1}$$

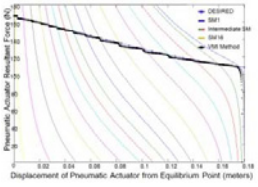


Musculoskeletal Model of Human Leg

M: Mass Matrix
C: Centrifugal Effect Matrix
V: Coriolis Effect Matrix
P: Pelvis Translation Matrix
G: Gravity Matrix
$M_{(H,K,A)}$: Joint Torques applied by Muscles
$\tau_{(H,K,A)}$: Joint Torques supplied by Pneumatic Actuators
$\theta_{(H,K,A)}$: Angle of (Hip, Knee, Ankle)
$A, B(\theta)$: Moment Arm Matrix for Muscles/Actuators

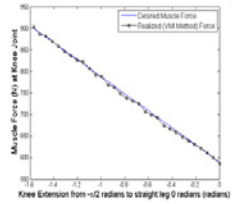
Variable Impedance Models to Obtain Desired Force-Profiles

A combined dynamic model of the pneumatic actuator (air spring) and lower leg was created to understand how the force of the pneumatic actuators will affect the dynamics of the human leg. The exoskeleton is designed as a tool to create and follow a variety of user-defined force-profiles to exercise different muscle groups. To achieve the desired force-profile, a simplistic method of variable mechanical impedances (VMI) is used over a more complicated force control design. This VMI Method involves different models or "states" to approximate a desired force-profile curve. Proportional valves then add/release air within the pneumatic actuator to achieve a certain force at a certain position.



VMI Method: Multiple States (SM) to approximate Desired Force-Profile; X-Axis: Displacement of Actuator, Y-Axis: Force

Yields



Desired Force-Profile (Blue) vs. Realized Force-Profile from VMI Method (Black); X-Axis, Knee Angle, Y-Axis: Force

Future Work

Controller:

A Controller is proposed to control the transitory states between one Impedance model to the next. Emphasis will be placed response time, accuracy to the desired force-profile and minimization of energy required to obtain force-profile

Experimental Prototype:

A prototype has been built to verify the Variable Impedance Model shown above. The setup involves:

•A Pneumatic Actuator,
•Two Pressure Sensors,
•Two 3-Way Proportional Valves,
• Air Pressure Supply,
•Air Tubing and a Chair to mount everything onto

References:

1.) http://1.bp.blogspot.com/_60A8zsWp2jY/TC088gcPDMI/AAAAAAAAAVI/Ft0ITxo-_I/s1600/2001orion.jpg
2.) http://www.nsbri.org/HumanPhysSpace/introduction/intro-bodychanges.html
3.) http://spaceflight.nasa.gov/gallery/images/station/crew-14/html/iss014e10591.html

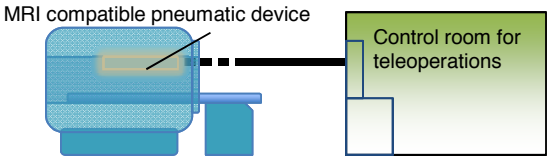
Pneumatically Operated MRI Compatible Haptic Interface

Bio-Inspired Robotics and Human Modeling Laboratory
Melih Turkseven, Ph.D. Student, Jun Ueda, Faculty Advisor
George W. Woodruff School of Mechanical Engineering

Introduction

Magnetic Resonance Imaging is capable of monitoring human soft tissue; hence, it is an excellent tool for diagnostics. Recent trend is to extend MRI's role on surgical operations and extensive brain researches.

Enabling simultaneous MRI usage during surgical operations will open a new era in medicine. That requires advanced robotic devices that can operate inside magnetic environment of MRI. Conventional robots interfere with MRI machine, but fluid powered robots with MRI-compatible haptic interfaces could be the solution.



Challenges

- Materials should be magnetically inert and non-conductive for safety
- Robots need to be compact
- Must create no artifact on the image

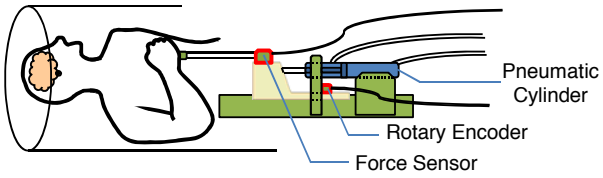
Goals

• Bring fluid power to MRI environment
• Create a multi-axes haptic interface which is compact and easy to integrate
• Enable accurate force feedback for robotic devices working in MRI environment

Collaborators

Haptic Interface

A pneumatically driven, active system is designed for advanced brain researches. As the subject pushes the end effector, this teleoperated interface deflects in a desired fashion so as to mimic a pre-selected surface condition. Same system could be used as a surgical robot, controlled from outside of MRI room.

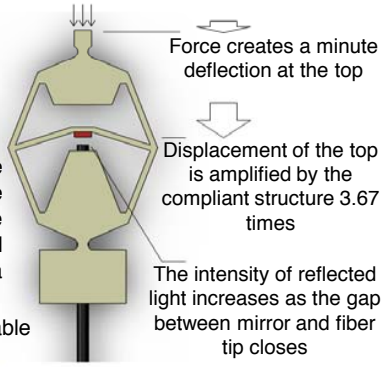


Force Sensor

A compliant structure amplifies the displacement in axial direction

Fiber optic cable carries light to the assembly. The intensity of reflected light is measured by a photodiode

Employs only 1 optic cable



Advantages:

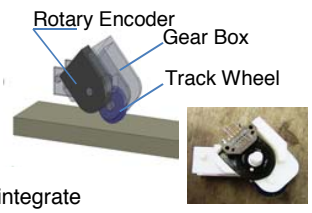
Compact, small-sized sensors
Minimal cable traffic, easy to install
Extendable to multi-axes
Clean and safe for MRI applications

Rotary Encoder

A commercial rotary encoder is connected to an external track wheel

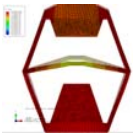
The wheel tracks the surface and encoder reads the rotation with a high resolution

Mobile, accurate and easy to integrate

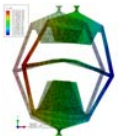


Progress

The compliant structure was optimized by a series of finite element simulations



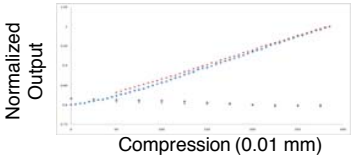
Axial displacement amplification is raised up to 3.67 times



Lateral disturbance does not change orientation of mirror to fiber tip

Experiments confirm finite element simulations

• The sensor decouples lateral disturbance



MRI Tests show that the sensor is invisible in MRI images

No sensor in the room | With sensor in the chamber | Difference (0.5%)

Future Work

Application of pneumatic line dynamics and close loop control with haptic feedback on the preliminary design

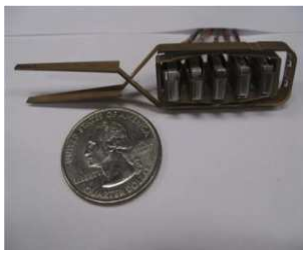
Extension to multiple-axes sensing and calibration

Robotics@GT & Intelligent Machines

Piezoelectric Self-Sensing Technique for Tweezer Style End-Effector

Bio-Robotics and Human Modeling Laboratory: Timothy McPherson, MS Student,
Dr. Jun Ueda, Faculty Advisor

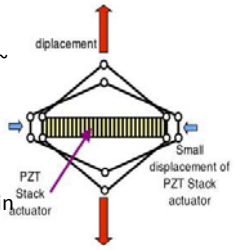
Piezoelectric Tweezers



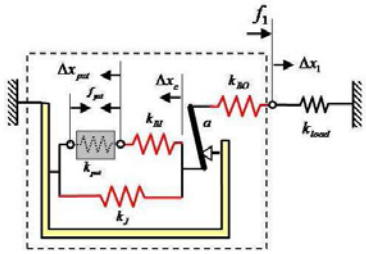
Previously, tweezers designed for robot assisted surgery were designed. They are driven by 5 commercially available Cedrat APA35xs actuators. They produce 9 mm of displacement at the tips if the tips are free or 1.1 N of force at tips if blocked. A piezoelectric material is electromechanically coupled, giving it usefulness as a sensor or an actuator. Previously one actuator was used solely as a sensor, wasting its actuation capability. The goal of this research was to develop a method to use one piezoelectric actuator as a sensor and actuator simultaneously → "self-sensing"

Strain Amplified Piezoelectric Actuators

- Piezoelectric ceramics generate high force (~200 N) but low strain (~ 10 µm)
- How to generate useable displacement?
- Rhombus strain amplification mechanism trades off force for displacement
- Tweezers consist of 3 layers of amplification



Lumped Parameter Model of Rhombus Strain Amplification Mechanism



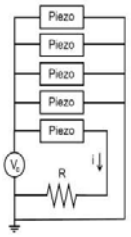
The lumped parameter model allows us to extrapolate the force and displacement at the tips from the force and displacement of the innermost piezoelectric actuator. At each successive level the force and displacement is a linear combination of the force and displacement of the level below it.

Piezoelectric Self-Sensing Technique

Based on piezoelectric constitutive equations, we can get expressions for the force and displacement of the innermost piezoelectric actuator based on the applied voltage and the induced charge. The self-sensing circuit is simply a shunt resistance in series with a piezoelectric actuator, which allows current (the first time derivative of charge) to be measured.

Loop law

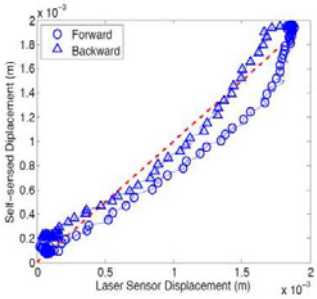
$$v_{pzt} = i\dot{R} - v_c \qquad f_{pzt} = \frac{-\dot{q}RC_{pzt} + q + v_c C_{pzt}}{d_{33}}$$

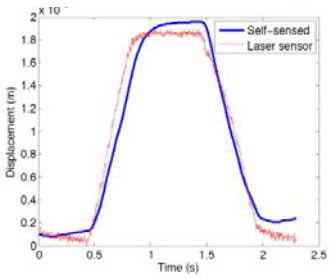


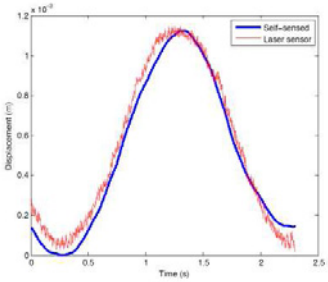
$$\Delta x_{pzt} = \frac{l}{y_{pzt} A d_{33}}(-\dot{q}RC_{pzt} + q + v_c C_{pzt}) + d_{33}(\dot{q}R - v_c)$$

Experimental Verification

Since force and displacement at the tips are linear combinations of force and displacement or the piezoelectric actuator, a multiple linear regression with respect to data from a laser position sensor is performed to find the correct coefficients. The maximum error was 20% of the dynamic range. The same calibration was then applied to a sinusoidal input, and showed 12% error.







Measurement of Human Muscle Stiffness & Its Application to Effective Human-Robot Interface Design

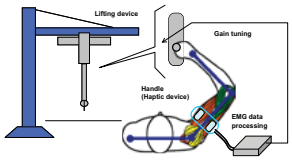
Billy Gallagher - PhD Student, Robotics
Timothy McPherson - MS Student, Mechanical Engineering
Jun Ueda - Faculty Advisor

Robotics@GT & Intelligent Machines

Bio-Robotics and Human Modeling Laboratory

Human Induced Haptic Instability

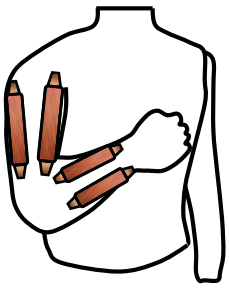
Haptics is a popular control method, as it has been found that touch is a very intuitive way for controlling a robotic device. Force feedback and haptic controllers are now common in areas from gaming to industrial machines. Haptic devices require physical contact between the operator and the machine, introducing feedback and creating a coupled system. This has been shown to introduce inherent instabilities due to the typical response of human operators. To attempt to correct for the negative effects of feedback, a control scheme is proposed that adjusts to changes in the way an operator's stiffness when working a haptic controller. Electromyogram (EMG) signals are expected to be a promising method of indicating stiffness.



Research Goals

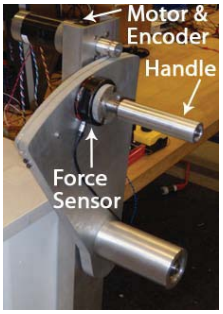
- ○Wearable biosignal (EMG) measurement device
- ○Method to determine stiffness from biosignals
- ○Control scheme that adjusts to stiffness level
- ○Investigate system performance and effectiveness

Muscle Cocontraction



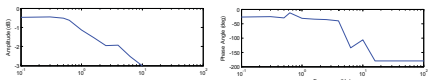
Joints are moved by at least two muscles that pull in opposite directions (antagonistic pair). An antagonistic pair contracting together (cocontraction) gives higher joint stiffness, since the result is no motion but more force on the joint. By detecting the cocontraction of antagonistic muscles it would be possible to estimate the stiffness of an operator's arm.

Hardware



1 DOF haptic Feedback Device

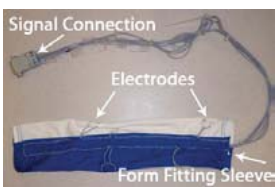
- ○Force Capability
  - •Max Force (Theoretical): 505.9 N
  - •Max Force (Practical): 100 N
- ○Frequency Response
  - •Max Frequency: 10 Hz



- ○Components
  - •Force/Torque Sensor
  - •Optical Encoder
  - •CompactRIO

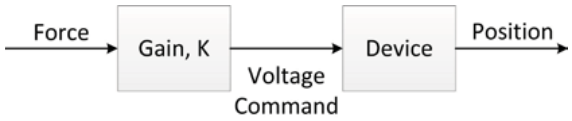
Wearable EMG Measurement Sleeve

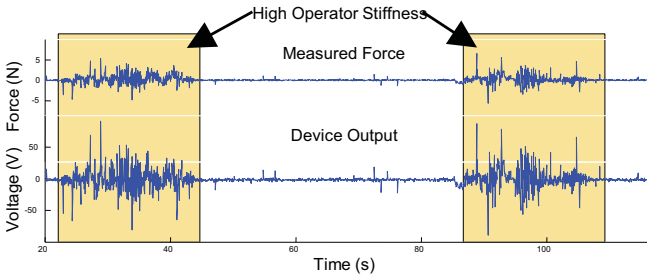
- ○Reusable EMG electrodes
- ○Easy to put on/take off
- ○Measures cocontraction
  - •Biceps/Triceps (elbow)
  - •Flexor/Extensor Ulnaris (wrist)



Force Assisting Instability

The haptic feedback device was used to reproduce instability with increased operator arm stiffness using a force assisting controller. The force applied to the device was scaled for the motor input to assist the operator. Increased operator arm stiffness resulted in more oscillation in the input force, creating oscillation in the device which grew with feedback, making it difficult for the operator to stabilize.

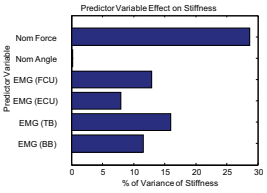


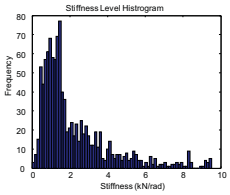


EMG/Stiffness Relationship

Experimental procedure*

- ○Participant holds handle
- ○Device pushes against participant
- ○Calculate arm end point stiffness from handle movement & force
- ○Vary position & device strength
- ○Regress stiffness on EMG signal





Statistically significant regression

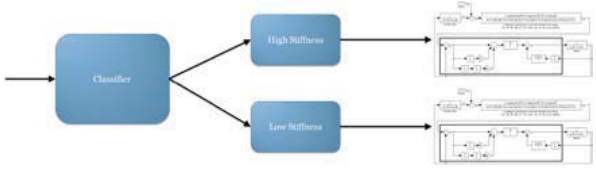
- ○Logarithmic relationship between EMG & stiffness
- ○EMG signals account for majority of variance of stiffness
- ○Position has little effect

Stiffness observations

- ○Participant arm stiffness varied widely
- ○Not necessarily minimum required for task

Future Work

- ○Design control system that adjusts based on stiffness
  - • Existing devices use impedance control
  - • Classify stiffness into ranges
  - • Switch between pretuned controllers appropriate for situation
- ○Evaluate system performance & effectives with human participant experiment



Life-long mapping of objects and places in domestic environments

John G. Rogers III

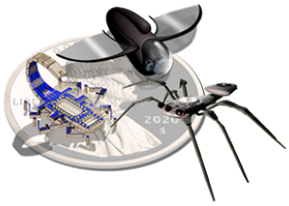
Advisor: Henrik I. Christensen

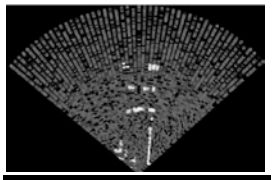
Georgia Institute of Technology

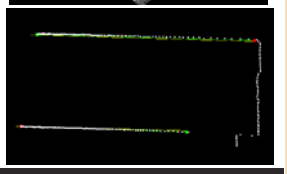
MAST

Objective: Perform enabling research and transition technology that will enhance warfighter's tactical situational awareness in urban and complex terrain.

- Data association in the presence of noise
- High level landmarks like walls and objects
- Synthetic aperture radar for long range, low power sensing
- Software toolkit for integration of multiple features





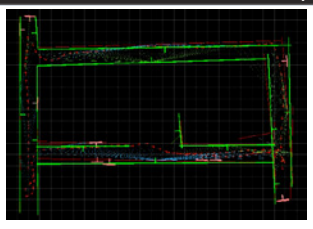


Learned object recognition, semantic data association

We developed a framework which can be taught to recognize visual landmarks.

- SVM based classification of Histogram of Oriented Gradient features
- Online image search to read text strings or receive additional semantic information about the object.
- Semantic information used for data association in mapping

MAST Joint Experiment



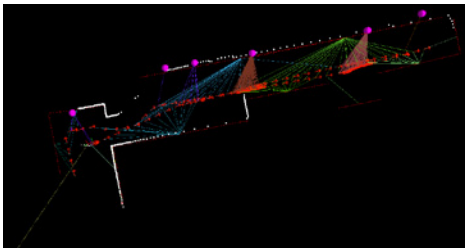
We performed a joint experiment for MAST in collaboration with **JPL** and the **University of Pennsylvania**. This experiment makes use of three robots which work together.

- Each of the robots used in this experiment is equipped with a laser scanner for measuring walls and a camera for locating doorways.
- Features are incorporated into a local map on each robot.
- Distributed-Data-Fusion algorithm used to combine local maps into a shared global map.
- An exploration strategy is used to efficiently cover the unknown environment.

GTmapper

We have developed a software toolkit which incorporates multiple types of sensor measurements in a graphical SLAM framework.

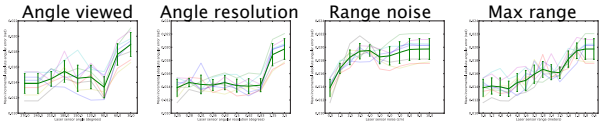
- Leverages GTsam library for full trajectory optimization
- Uses Joint Compatibility for robust data association in the presence of noise
- Easily adapted for new sensor types and landmarks with M-Space representation



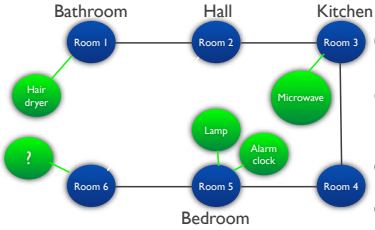
Sensory precision for SLAM

We explored the relationship between sensory accuracy and SLAM performance by varying parameters such as:

- maximum range
- noise
- angular precision
- viewable angle
- Also compared popular sensors (URG, UTM30, and LMS291).



Life-long mapping of objects in domestic environments



Conditional Random Field model for joint object and place recognition

- Room adjacency compatibility functions
- Object in room compatibility functions

Technical Foundations

- GTMapper
- SURF matching for obj rec
- Exploration controller
- 3D surface extraction
- 3D object extraction
- GoogleGoggles interface

Future Work

- CRF model
- Pose factors in GTMapper
- Bag of visual words model to accelerate object recognition
- Social networking interface

Supported by the Army Research Lab

Social Media Enabled Robotics/ Push Planning for Object Placement in Clutter

Ph.D Student: Victor Emeli
Advisor: Henrik Christensen
The Georgia Institute of Technology

Enhancing the Robotic Service Experience Through Social Media

The goal of this project is to leverage the attributes of social media to improve the performance and acceptance of robotic service in a restaurant or service environment. The long term goal of this research is to develop techniques that allow a robot to enhance its human robot relational capabilities through accessing information residing in various social media technologies. These techniques will allow a robot to learn individual preferences, habits, and patterns, which will lead to more personal, relevant, and longer term Interactions.

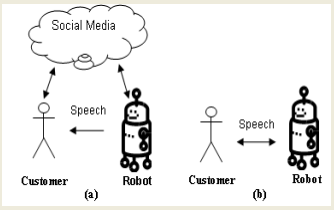
- Robot Server delivers service to customer using Voice or Social Media.
- Social Media Enabled Robot utilizes Facebook and Twitter.

Social Media Enabled Robot Server

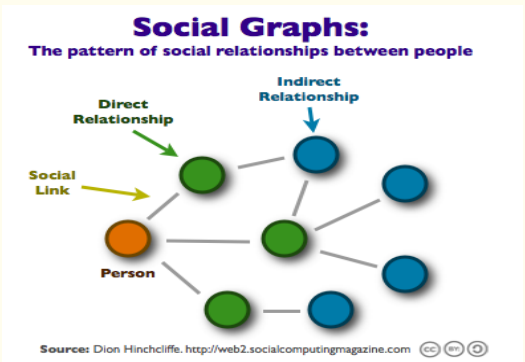
Two-Group Experiment: Voice vs. Social Media Enabled Robot

These hypotheses were tested by conducting an online 2 x 2 factorial design scenario surveys with counterbalancing and randomization of scenarios. The independent variables are the service methods (social media and voice) and the service communication (difficulty and no difficulty).

- Submitted to Ro-MAN 2011
- Likert Scale Survey administered to 59 participants that watched videos of the interaction.
- Robot that utilized both Facebook and Twitter obtained highest rating.



Computational Models for Analyzing Social Media



This research is a first step to developing computational models that will allow a robot to engage in meaningful longer term interactions with humans. The ability to access up-to-date social information will allow the robot to learn individual preferences, patterns, and engage in more relevant conversations.

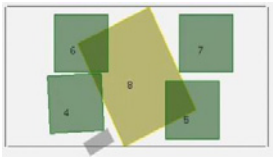
Dynamic Push Planning for Object Placement in Clutter

PR-2 analyzes table top to plan for object placement

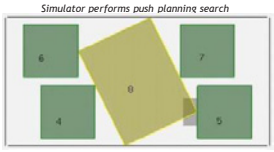
A novel planning algorithm for the problem of placing objects on a cluttered surface such as a table, counter or floor. When no continuous space is large enough for direct placement, the planner leverages means-end analysis and dynamic simulation to quickly find a sequence of linear pushes that clears the necessary space..

- Algorithm submitted to IROS 2011
- Simulator modified for execution on the PR-2

Iterative Breadth First Search Algorithm w/Goal Checking



Simulator performs push planning search



The Iterative Deepening Breadth First search (IDBFS) algorithm attempts to clear the placement object footprint (green).

The state space inherently has many shallow solutions. To take advantage of this, IDBFS searches for multiple Level 1 goals in the search tree before moving to the next level.

A goal check algorithm checks if a subgoal is reached for every push attempted below Level 1 in the search tree.

Open Loop vs. Closed Loop Control

The implementation on the PR-2 compares open-loop against closed loop control using dynamic replanning. A lower bound for the number of linear pushes that are consistent with the results of the simulator is empirically investigated. This lower bound will indicate the number of consecutive pushes that can be reliable attempted before removing uncertainty through replanning is necessary.

PR-2 executes final plan in the real world

Once the simulator successfully implements a plan, it is transferred to the PR-2 for execution. Augmented Reality tags are used to simplify the identification of the table and clutter objects.

Future work will include generalizing performance to arbitrary objects through 3D point cloud object segmentation.

SEMANTIC SLAM
ROBOTS NAVIGATING LIKE HUMANS

CARLOS NIETO-GRANDA & HENRIK CHRISTENSEN
College of Computing, Georgia Institute of Technology
http://www.cc.gatech.edu/~cnieto6
carlos.nieto@gatech.edu

Motivation

SEMANTIC SLAM

The relationship between a place and the knowledge that is associated with specific tasks e.g.(functionality, objective location).

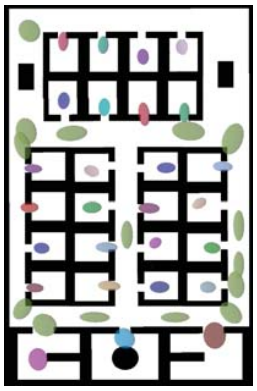
- Algorithms for recognition and classification of spaces into semantic regions.
- The association of semantic labels with spatial regions is based on human guide or place categorization.
- Object recognition allows the robot to interact with the world for avoiding obstacles and doing specific tasks.
- The robot can navigate in the environment with its own semantic map partition.

Our mobile robot "Jeeves" navigating in the Aware Home Lab. with and Ideal place and object categorization.

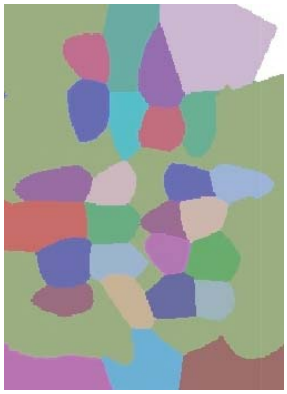
The main goal of this research project is to provide methods for automated recognition and classification of spaces in both indoor and outdoor environments into semantic regions.

Map Partitioning using Regional Analysis

The robot is able to navigate the environment by using a map based on the multivariate probability distributions for a set of semantic labels.



a) A visualization of a simulated office environment. the Gaussians in our model and different colors represent different spaces.



b) A visualization of the decision Colored ellipses represent boundaries of the regions.

**These visualizations looks better in color.*

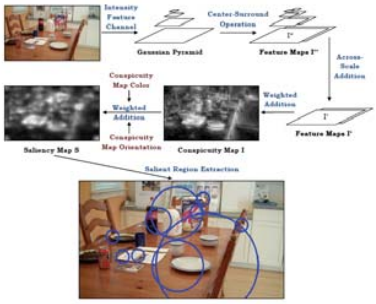
3D Object Recognition for Semantic Labeling Applications

In the real world a robot needs to be able to learn about the environment without a human teacher.

High-level features can give a robot a better understanding of the environment, and allow it to recognize environment-specific tasks such as grasping a cup.

A visualization of the process to segment and extract objects from a point cloud.

Automated Place Categorization



Our approach is an automated visual place categorization.

The idea is to give the robot the capability to identify which room of the building it is in without human guidance.

The mechanism for detecting objects is the same as how a human observer would categorize a new location.

A visualization of the visual attention mechanism to find and detect interesting objects.

Service Robots in Home Environments

In order to evaluate our methods, we participate in the international annual competition for autonomous service robots: "RoboCup".

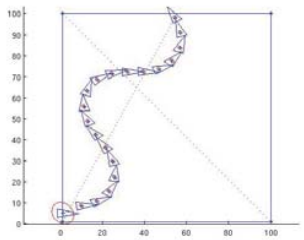
The competition comprises of a set of benchmark tests for service robot performance in a realistic non-standardized home environment setting.

GT@Home performing a object recognition and grasping task.

SLAM in Industrial Environments

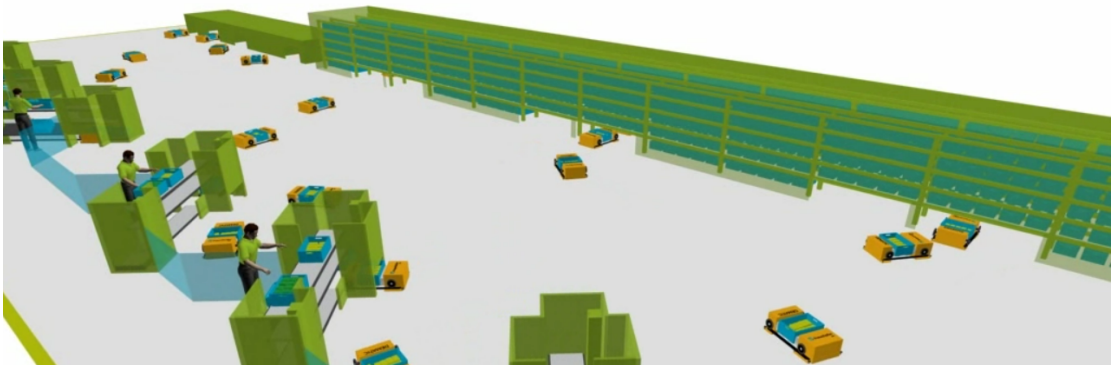
Service robots are not only in home environments, we can find them in industrial environments too.

Our methods provide the portability to run in an industrial environment such as localizing a robot in a warehouse which the robot needs to navigate in a dynamic environment.



PLANNING IN WAREHOUSE MANAGEMENT SYSTEMS

Pushkar Kolhe and Henrik I. Christensen
Robotics & Intelligent Machines, College of Computing
Georgia Institute of Technology
{pushkar, hic}@cc.gatech.edu



BENEFITS OF AUTOMATED WAREHOUSES

- Efficient use of storage space can save 20% costs
- Automated distribution centers can cut labor cost by 10-20%
- Efficient mixed palletizing can reduce transportation costs by 4-8%
- Orders can be fulfilled faster, with less errors
- Automated Warehouses are **scalable**
- Automated Warehouses can **re-organize** themselves

PROBLEM FORMULATION

Warehouse Management Systems (WMS) includes **automated storage and efficient retrieval of goods**, **navigation planning for multiple AGV's**, **mixed palletizing**, and **shipping in reverse unloading sequence**.

For efficient retrieval for SKU's, they are distributed throughout the warehouse in mixed pallets. For shipping, SKU's are closely packed in mixed pallets. We allow our model warehouse to re-organize itself by using a palletizing de-palletizing cycle using a buffered storage space. We formalize this known NP-hard problem as a constrained optimization problem. We try to optimize:

- Minimize delay in shipping
- Maximize storing and packing density

Our constraints include:

- The available storage space
- Sequence in which the orders have to be delivered

A KNAPSACK VARIANT

We formulate the constrained optimization problem as a **multiple knapsack problem** (MKP). The incoming orders can be modeled as a **stochastic process** to find distribution of SKU's within the warehouse for efficient retrieval. Exact algorithms are computationally inefficient to solve this problem. We explore approximate randomized algorithms which can give probabilistically complete solutions in polynomial time.

$$\max \sum_{i=1}^{m} \sum_{j=1}^{n} p_j x_{ij} \tag{1}$$

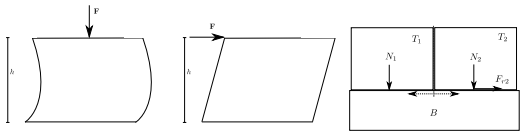
subject to

$$\sum_{j=1}^{n} w_j x_{ij} \leq c_i, \; i = 1, \ldots, m, \tag{2}$$

$$x_{ij} \in \{0, 1\}, \; i = 1, \ldots, m, j = 1, \ldots, n. \tag{3}$$

BENCHMARKING

We study benchmarks and metrics for determining which planners work better in a WMS. For a warehouse, we are mainly interested in an **asymptotic upper bound** for the running time to illustrate the order of magnitude of the increase in running time when the number of orders is doubled.
We are also interested in the amount of storage space used. SKU's are stored in mixed pallets. For mixed pallets we are interested in their **density** and **stability** of a pallet. We model pallets using *finite element modeling principles* and establish new metrics.



VMAC COMPETITION

The Virtual Manufacturing and Automation competition is held annually at ICRA to highlight WMS systems designed by students. The planning systems are evaluated with new evaluation software that use new benchmarks.

- **PalletViewer** is used to compare pallet density.
- **USARSim** can model pallet stability by using physics.
- **KUKA Robot Cell** test on a real robot cell.

FUTURE WORK

- Provide a framework of benchmarks and metrics that will allow vendors to compare their planners.
- Provide a general framework of planning for different kinds of automated warehouses.
- Construct a mini warehouse at NARC to test our algorithms. The warehouse will be able to model automated storage and retrieval, mixed palletizing and depalletizing and re-organizing using buffered storage space.

Planar Surface Mapping

Alexander J. B. Trevor, John G. Rogers III, Henrik I. Christensen
Robotics & Intelligent Machines, College of Computing
Georgia Institute of Technology
{atrevor, jgrogers, hic}@cc.gatech.edu

Introduction

We present a feature based mapping technique that allows for the use of planar surfaces such as walls, tables, counters, or other surfaces as landmarks. These planar surfaces are detected in 3D point clouds, and provide measurements via their surface normal and perpendicular distance. We also map the convex hulls of the observed planar patches and use these for data association, allowing multiple non-overlapping coplanar landmarks to exist in the map. Maps of such planar surfaces could be useful for semantic mapping, and could benefit mobile manipulation tasks.

Robot Platform & Sensor Data

The robot used in this work is equipped with a Hokuyo UTM-30-LX laser range finder mounted on a Directed Perception DP-46-70 pan tilt unit. Tilting the 2D laser scanner allows 3D point cloud data to be collected, by using the sensor location and tilt angle to project the points into 3D.

The robot platform used in this work is shown above. The 3D laser sensor can be seen at the top of the robot.

Plane Detection

3D point clouds collected throughout the area to be mapped are processed in order to detect planar surfaces. To do this, we use the well known RANdom SAmple Consensus (RANSAC) method for model fitting. In our case, we are fitting planes to the full point cloud to determine the largest plane present in each cloud. We then remove the inliers associated with this plane from the point cloud, and repeat the process in order to detect additional planes. Convex hulls are then calculated for detected planar regions.

For much of our point cloud processing, we use the Point Cloud Library (PCL) developed by Rusu and others at Willow Garage, which includes a variety of tools for working with 3D point cloud data including RANSAC plane fitting, outlier removal, and euclidean clustering methods.

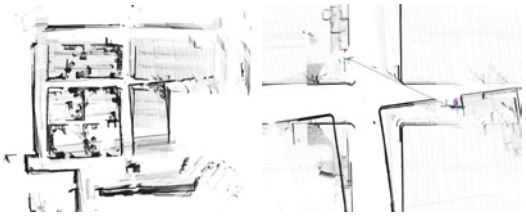
Plane Mapping

Our SLAM system uses the GTSAM library developed by Dellaert [1]. GTSAM approaches the graph SLAM problem by using a factor graph that relates landmarks to robot poses through factors. The factors are nonlinear measurements produced by measuring planar surfaces detected in point cloud data. While planar surfaces are mapped in 3D, note that we constrain our robot trajectory to the 2D groundplane, so poses consist of (x, y, θ).

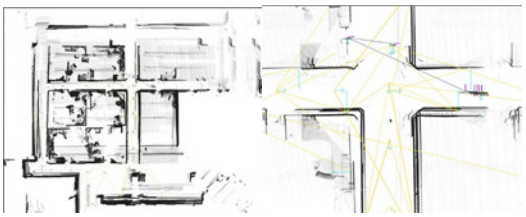
Using the planar surface measurement function and its associated Jacobians, we can utilize planar normals and perpendicular distances as landmarks in our SLAM system. During optimization, the landmark poses and robot trajectory are optimized. After optimization, the resulting map consists of planar surfaces as represented by their surface normals and convex hulls in the map frame, along with the optimized robot trajectory.

Mapping Results

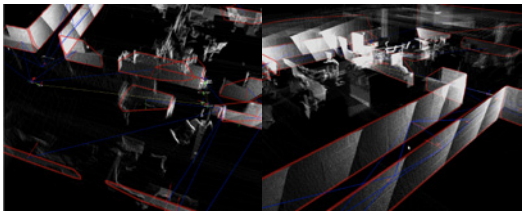
The mapping system was tested on data collected at the Georgia Institute of Technology's College of Computing building. Mapping results are presented in the figures below.



Shown here are pointclouds plotted in the odometric coordinate frame without optimization. An overview is shown on the left, with closeup on the right. Note the significant error in alignment.



Pointclouds plotted in the map frame, after optimization. An overview is shown on the left, with closeup on the right. Note that the alignment has been much improved.



The convex hulls of the mapped planar regions can be seen in the above figures, showing the location and extent of the mapped surfaces.

Conclusion

- Point cloud data can be processed to detect planar surfaces along with their convex hulls.
- Planar surfaces can be mapped using their surface normals and perpendicular distances, along with their convex hulls.
- The resulting maps include information on both the locations and extent of surfaces, which could be useful for semantic mapping and mobile manipulation tasks.
- Mapping results were presented for an office environment demonstrating the system's ability to loopclose.

References

[1] F. Dellaert, M. Kaess Square Root SAM: Simultaneous Localization and Mapping via Square Root Information Smoothing *IJRR*, Vol. 25 (12), 2006.

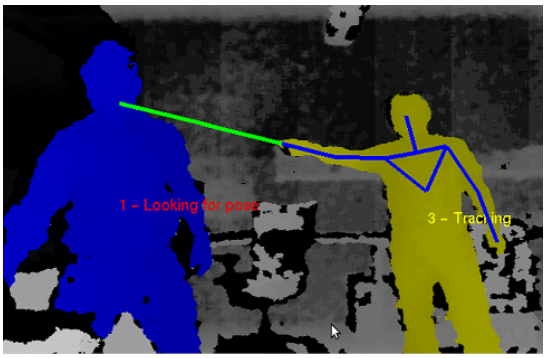
Person Following by using Natural Gestures for User Selection

Akansel Cosgun and Henrik I. Christensen
Robotics & Intelligent Machines, College of Computing
Georgia Institute of Technology
akanselcosgun@gatech.edu, hic@cc.gatech.edu

Abstract

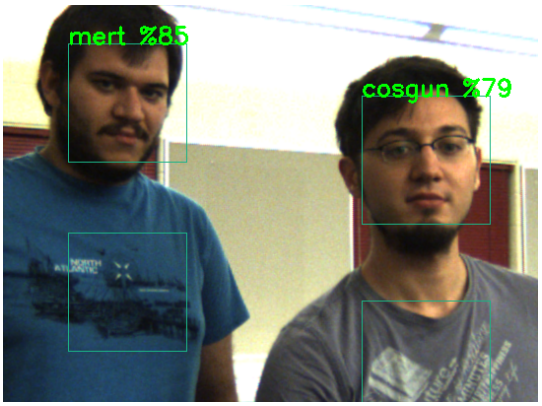
We present a system that is able to follow a specific person by using leg tracking. The user selection is done by using natural gestures such as speech commands or pointing to someone else. The system is able to learn the facial features/shirt color of people and tell who is who.

User Selection



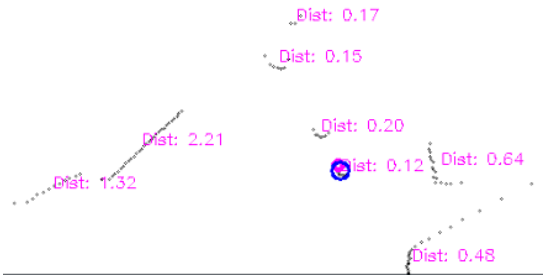
- User selection is done by skeleton tracking, using the depth image from a Kinect sensor.
- User can say 'Follow me' to start following behavior, or alternatively point to another person and say 'Follow him/her'.
- Pointing gesture is recognized by checking if the direction of the forearms stays stationary for a specified period of time.
- The 'pointed' person is the closest person to the ray emaneting from the hand of the user.

Person Recognition



- Real-time on-line training on faces/shirts for identification of people.
- Faces are detected by Viola-Jones face detector of OpenCV. Facial features are learnt by applying Principal Component Analysis(PCA) on extracted eigenfaces.
- Shirt histogram is learnt in normalized RGB domain and is adaptively updated.
- When a new person is detected, both face recognition and shirt color matching contributes to scoring.
- A database of people is kept as the robot meets new people.

Leg Tracking



- Used a SICK laser scanner at 30cm height to detect leg hypotheses.
- In 15000 sample leg scans, geometric parameters of legs such as Width, Linearity, Concavity, Circularity are calculated. The mean and standard deviation of those parameters are used in leg detection.
- Mahalanobis distance to feature vector for every segment in the laser scan are calculated on the run. Lower distance values are better leg hypotheses.
- After the initial detection, leg parameters are adaptively updated specific to the user's legs.
- The user's legs are very efficiently tracked by a Kalman Filter.
- The tracker is scalable to tracking different objects since the geometric parameters are updated on the run.

Following

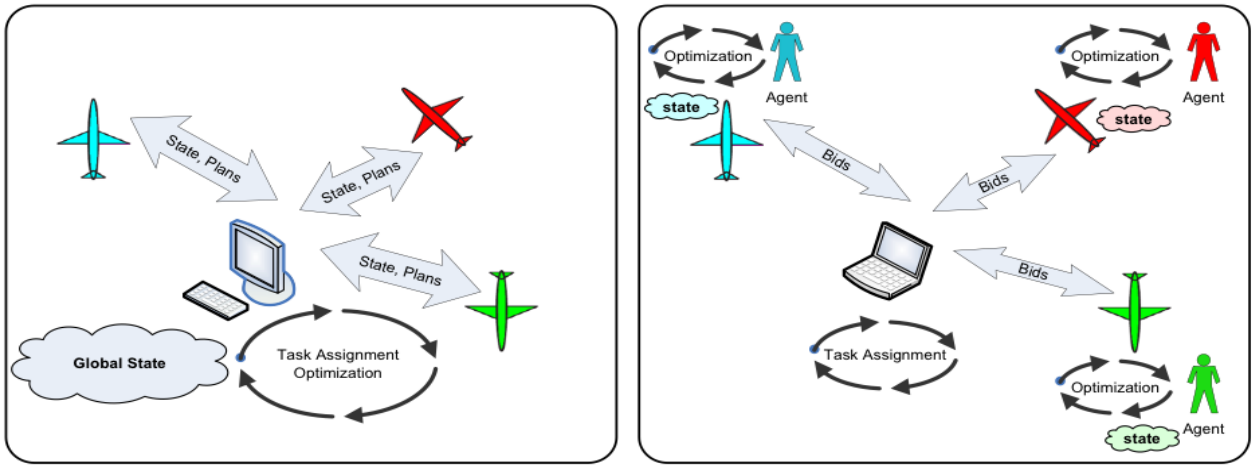
- Implemented the system on a Segway based robot with caster wheels.
- Software architecture is Robot Operating System(ROS) running on a Linux machine. Coded in C++.
- Robot tries to stay 0.8m away from the user. It avoids obstacles and other people nearby.
- Goal position for robot is updated at 10Hz, and a path is planned in by searching in the grid map.
- Robust to temporary occlusions. Even if other people pass between the robot and the user, it still follows the user.
- Robot is able to go through narrow passages such as doorways.

Application Areas

The system can be used in areas such as giving a home tour to robot and annotating people/places by pointing, carrying packages of people in a supermarket and following workers in a factory environment to carry heavy tools.

Multi-Agent Auctions for Task Allocation

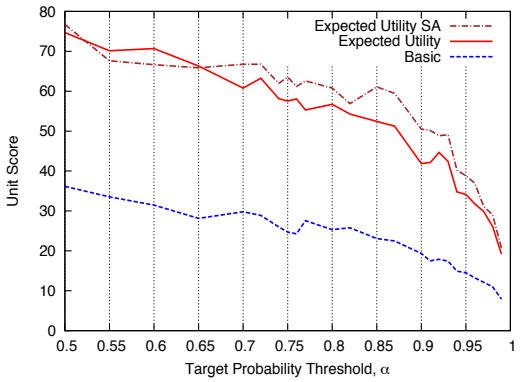
Charles Pippin, pippin@gatech.edu; Advisor: Henrik Christensen

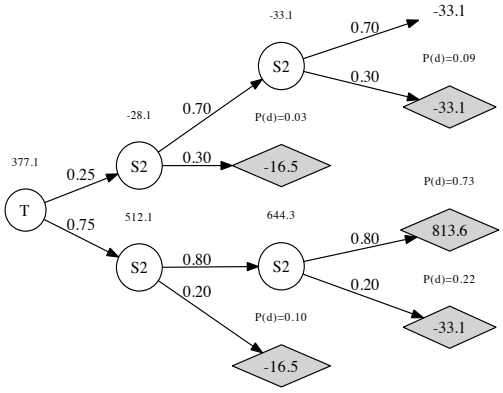


A Centralized vs. Decentralized approach. The centralized approach (left) performs task optimization for all assets using centralized computation. This requires overhead of state and plan communication between all assets and the central node. While central computations can lead to bottlenecks and long run-times, decentralized approaches can run much faster because computation can be performed in parallel. A decentralized, auction-based method (right) uses distributed agents to analyze new tasks. Agents can be distributed logically and physically in the system. Communication requirements are also lower, sending only bid and assignment information.

This research focuses on algorithms in which heterogeneous agents can distribute tasks among themselves in a decentralized way.

Areas of research include incorporating different sensor characteristics and agent reliability into multi-agent auctions, and learning bidding behavior.





Reference: Charles Pippin and Henrik Christensen. *A Bayesian Formulation for Auction Based Task Allocation in Heterogeneous, Multi-Agent Teams*. SPIE Defense Security and Sensing, April 2011, Orlando

Applying expected utility to bid valuation in an auction can be used to combine agents with different sensor qualities.

Mobile Manipulation for Service Robots in Domestic Environments

Jacob Huckaby
Robotics and Intelligent Machines
Henrik I. Christensen, Advisor

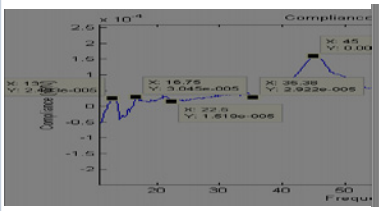
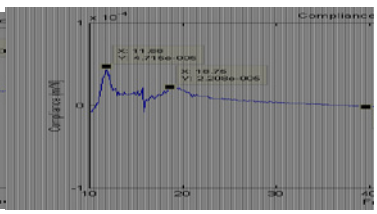
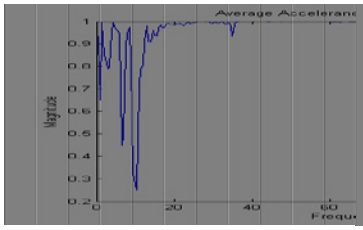
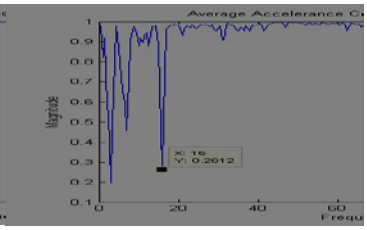
Dynamic Characterization of Light-Weight Robots

High precision tasks are an important part of the manufacturing industry. For example, safety constraints require that some manufacturing must be done to a very high degree of accuracy. With the recent availability of smaller, light-weight robot manipulators, the question arises of whether it would be feasible to use them in tasks that would otherwise be unsuitable for the standard large industrial robots. Possible tasks would include those in tight or constrained spaces, such as precision drilling. For this to be possible, the light weight robot would need to be stiff enough to be able to meet safety constraints.

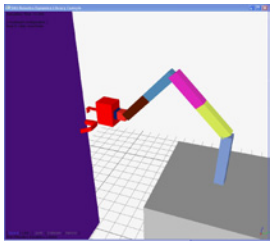
To that end, this study analyzed the dynamic characteristics of small, light-weight manipulators to determine whether it would be possible for them to be used in manufacturing tasks. Two light-weight robot manipulators were considered in the study: the KUKA Light-Weight Robot (LWR) and the KUKA KR5 Sixx.

- Experimental modal analysis used to determine dynamic parameters
- Two characteristics were compared
  - Coherence: $$C_{xy}(\omega) = \frac{|P_{xy}(\omega)|^2}{P_{xx}(\omega)P_{yy}(\omega)}$$
  - Compliance Frequency Response Function: $$\alpha(\omega) = \frac{X(\omega)}{F(\omega)}$$
- Identified end effector resonant frequencies and compliance
- LWR more stiff than KR5 Sixx across 1-100 Hz frequency range



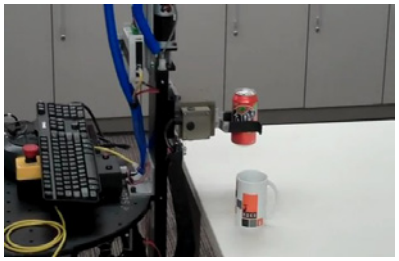
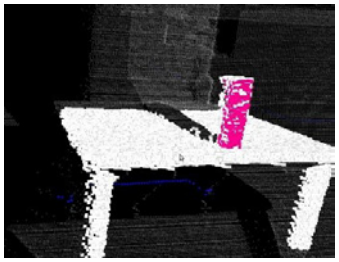
Door Opening Strategies



This project focused on the design and implementation of a system that would be able to learn how open unknown doors based on experience opening doors with different types of door handles. We were able to implement several algorithms that allowed the robot to open a previously unseen door based on initial door pose information, and correct for errors in initial guesses using dynamic reclassification to attempt different opening approaches.

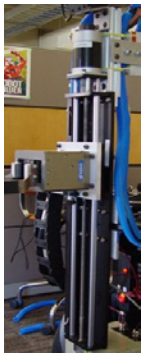
- Initial testing was done in srLib, a simulation library for multi-body dynamics
- Force-torque information from the joints was used to classify door opening strategies

Low Degree-of-Freedom Manipulation

This work explored the use of a low degree-of-freedom manipulator as a viable and cost-effective solution to the mobile manipulation problem in a home setting. The project utilized the Jeeves mobile manipulation platform. Through experimental results we have shown that with a single degree of freedom manipulator and the control scheme proposed and implemented on the Jeeves platform, we can achieve reliable and repeatable results for a number of varied household object types. Perception was performed using a 3D laser scanner to create a 3D point cloud, from which table top objects could be segmented. Ten objects were selected to serve as the representative categories of general objects found in domestic environments.

- Objects are classified by shared geometric properties
- Controller works well for simple geometries (cylinders, boxes)
- Failures occurred most often with odd-shaped objects at the corner of the grasping surface




Object	Success (%)	Shape	Overall (%)
Pringles	100	narrow	100
Fanta	100		
Coke	100		
Cable Ties	91.7	wide	93.75
Clorox	95.8	wide	
Mango Tea	100	box	91.67
Earl Grey	83.3	box	
Goo Gone	91.7	odd	83.33
Vitamins	100	odd	
Buzz	58.3	odd	

Localization, Navigation and Mapping for single and multi-robot systems

Manohar Paluri, Henrik Christensen and Frank Dellaert
Robotics & Intelligent Machines(RIM) lab, Gatech

Multi-Robot SLAM

Figure 1a. Scarab robot

Figure 1b. Fleet of scarabs used to demonstrate multi-robot SLAM using wall features and door features in a distributed data fusion framework.

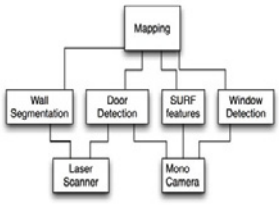


Figure 1c. Architecture of the system

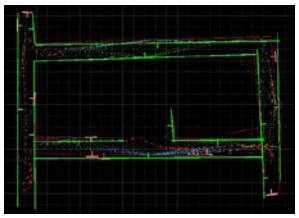


Figure 1d. Optimized map explored by three robots.

Abstract

The system developed makes use of three robots which work together to explore and map an unknown environment. Each of the robots used in this experiment is equipped with a laser scanner for measuring walls and a camera for locating doorways. Information from both of these types of structures is concurrently incorporated into each robot's local map using a graph based SLAM technique. A Distributed-Data-Fusion algorithm is used to efficiently combine local maps from each robot into a shared global map. Each robot computes a compressed local feature map and transmits it to neighboring robots, which allows each robot to merge its map with the maps of its neighbors. Each robot caches the compressed maps from its neighbors, allowing it to maintain a coherent map with a common frame of reference.

High Fidelity Localization & Navigation

Figure 2a. SICK Laser scanner

Figure 2b. Segway robot with two SICK scanners

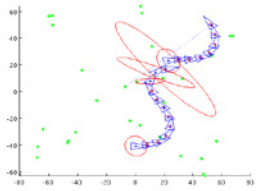


Figure 2c. EKF based Localization

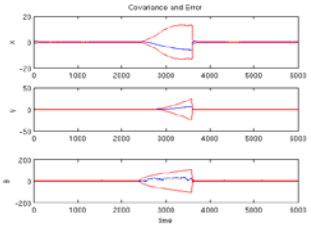


Figure 2d. Covariance and estimate error over time.

Abstract

We present a solution for autonomous localization and navigation of a huge mobile platform(Kuka's omni-move platform) using laser scanners. The proposed system will be tested and used for autonomous painting of aircrafts for Boeing in their premises. The focus of the current work lies in using two laser scanners for high precision and high accurate localization and smooth navigation in the cluttered painting hangar. Multiple omni-move platforms will be using our algorithm to move in the paint hanger and automatically carry out the required tasks. This environment is hostile for visual sensors because of low visibility, sparse/no visually features etc. and floating paint particles. Such constraints lead us to use Laser scanners. Our system uses the walls of the room as features for localizing. We use range weighted Hough transform for localization and RRT* planning algorithm for planning. We tested our approach on a segway robot and share the visuals and performance results on this dataset. Due to confidentiality reasons we only share the performance metrics on actual paint hangar data. Our system meets the real-time and high accuracy requirements consistently on multiple runs.

Similar shaped Object detection & discrimination

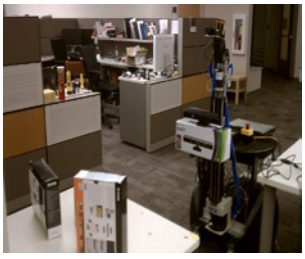


Figure 3a. Kinect setup on the Segway with objects on the table

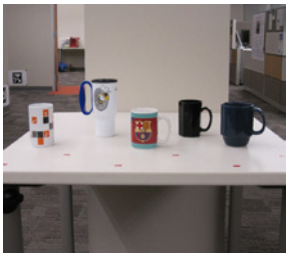


Figure 3b. Various cups that are similar shaped

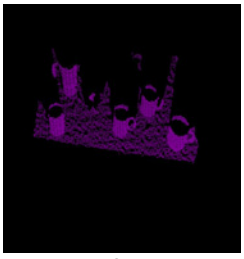


Figure 3c. Sample point cloud from Kinect sensor of the objects on the table

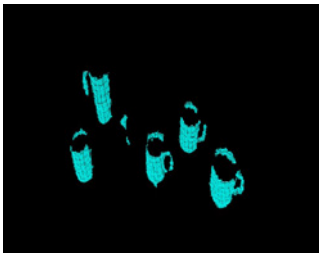


Figure 3d. Segmentation of objects

Abstract

The holy grail of computer vision is to solve the problem to identify and recognize objects without a previous labeling for example using artificial markers to recognize the object fast in the environment. Most of the objects are different in shape, color and size. However some of them are not. In this work we present a technique that allows a robot to find objects on a table, classifying it by shape and labeling as a different objects without non-previous knowledge of it. All the experiments were done using real data acquired with our robot platform and Kinect and laser sensors.

Robotics@GT & Intelligent Machines

Robust 3D Visual Tracking Using Particle Filtering on the SE(3) Group

Changhyun Choi and Henrik I. Christensen
Robotics & Intelligent Machines, College of Computing
Georgia Institute of Technology
{cchoi, hic}@cc.gatech.edu

Motivation

We present a 3D model-based visual tracking approach using edge and keypoint features in a particle filtering framework. While most of the edge-based tracking has made an assumption that an initial pose is given, we employ keypoint features for initialization of the filter.

Contributions

- We employ keypoint features to initialize particles to highly probable states based on pose estimates from keypoint correspondences.
- For better accuracy, we refine edge correspondences between the projected model edges and the image edges via a RANSAC.
- Instead of random walk models, we apply a first-order autoregressive (AR) state dynamics to guide particles more effectively.
- Our approach monitors the number of effective particles and use the value to decide when the tracker requires re-initialization.

Particle Filter

The discrete system equations on the $SE(3)$ group is formulated based on [1] as follows:

$$\mathbf{X}_t = \mathbf{X}_{t-1} \cdot \exp(A(\mathbf{X}, t)\Delta t + \mathbf{dW}_t\sqrt{\Delta t}), \tag{1}$$

$$d\mathbf{W}_t = \sum_{i=1}^{6} \epsilon_{t,i}\mathbf{E}_i,\ \boldsymbol{\epsilon}_t = (\epsilon_{t,1}, \ldots, \epsilon_{t,6})^{\mathsf{T}} \sim \mathcal{N}(\mathbf{0}_{6\times 1}, \boldsymbol{\Sigma}_w) \tag{2}$$

where $\mathbf{X}_t \in SE(3)$ is the state at time t, $A : SE(3) \mapsto se(3)$ is a possibly nonlinear map, $\mathbf{dW}_t$ represents the Wiener process noise on $se(3)$ with a covariance $\boldsymbol{\Sigma}_w \in \Re^{6\times 6}$, $\mathbf{E}_i$ are the i^{th} basis elements of $se(3)$.
The corresponding measurement equation is then:

$$\mathbf{Z}_t = g(\mathbf{X}_t) + \mathbf{n}_t,\ \mathbf{n}_t \sim \mathcal{N}(\mathbf{0}_{\mathsf{N_Z}\times 1}, \boldsymbol{\Sigma}_n) \tag{3}$$

where $g : SE(3) \mapsto \Re^{\mathsf{N_Z}}$ is a nonlinear measurement function and $\mathbf{n}_t$ is a Gaussian noise with a covariance $\boldsymbol{\Sigma}_n \in \Re^{\mathsf{N_Z}\times\mathsf{N_Z}}$.
The first-order AR state dynamics on the Lie group is applied for more effective particle propagation as:

$$\mathbf{X}_t = \mathbf{X}_{t-1} \cdot \exp(\mathbf{A}_{t-1} + \mathbf{dW}_t\sqrt{\Delta t}), \tag{4}$$

$$\mathbf{A}_{t-1} = \lambda_a \log(\mathbf{X}_{t-2}^{-1}\mathbf{X}_{t-1}). \tag{5}$$

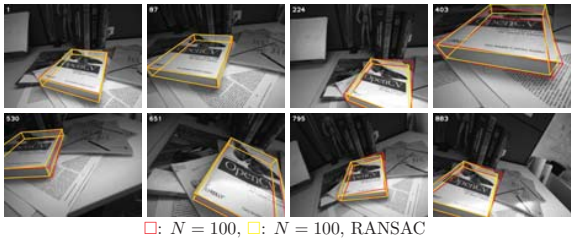
Algorithm

Data: $\mathcal{I} = \{\mathcal{I}_0, \mathcal{I}_1, \cdots, \mathcal{I}_I\}, \mathcal{F} = \{(\mathbf{p}_1, \mathbf{P}_1), \cdots, (\mathbf{p}_F, \mathbf{P}_F)\}$
Result: $\mathcal{S} = \{\mathcal{S}_0, \mathcal{S}_1, \cdots, \mathcal{S}_I\}$
Params: $\boldsymbol{\Sigma}_w, \lambda_a, \lambda_v, \lambda_e, N_{thres}$

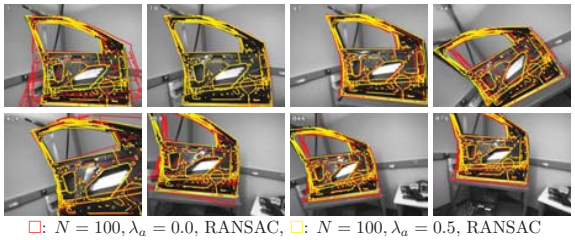
$t \leftarrow 0$; $init \leftarrow 1$; $A_0 \leftarrow 0_{4\times 4}$
while $\mathcal{I}_t \neq 0$ **do**
  if $init = 1$ **then**
    $\mathcal{S}_t \leftarrow$ `InitParticle`$(\mathcal{I}_t, \mathcal{F})$
    if $\mathcal{S}_t \neq \{\phi\}$ **then**
      $init \leftarrow 0$
  else
    for $n \leftarrow 1$ **to** N **do**
      $\mathbf{X}_t^{*(n)} \leftarrow$ `Propagate`$(\mathbf{X}_{t-1}^{(n)}, \mathbf{A}_{t-1}^{(n)}, \boldsymbol{\Sigma}_w)$
      $\mathbf{A}_t^{*(n)} \leftarrow$ `AR_vel`$(\mathbf{X}_t^{*(n)}, \mathbf{X}_{t-1}^{*(n)}, \lambda_a)$
      $\mathbf{Z}_t^{*(n)} \leftarrow$ `Measurement`$(\mathbf{X}_t^{*(n)}, \mathcal{I}_t)$
      $\dot{\mathbf{Z}}_t^{*(n)} \leftarrow$ `RANSAC`$(\mathbf{Z}_t^{*(n)})$
      $\tilde{\pi}_t^{*(n)} \leftarrow$ `EdgeLikelihood`$(\dot{\mathbf{Z}}_t^{*(n)}, \lambda_v, \lambda_e)$
      $\dot{\mathbf{X}}_t^{*(n)} \leftarrow$ `IRLS`$(\mathbf{X}_t^{*(n)}, \dot{\mathbf{Z}}_t^{*(n)})$
    $\pi_t^* \leftarrow$ `Normalize`$(\tilde{\pi}_t^*)$
    $\widehat{N_{eff}} \leftarrow$ `Neff`(π_t^*)
    if $\widehat{N_{eff}} \geq N_{thres}$ **then**
      $\mathcal{S}_t \leftarrow$ `Resampling`$(\mathcal{S}_t^*)$
    else
      $init \leftarrow 1$
  $t \leftarrow t + 1$

Experimental Results

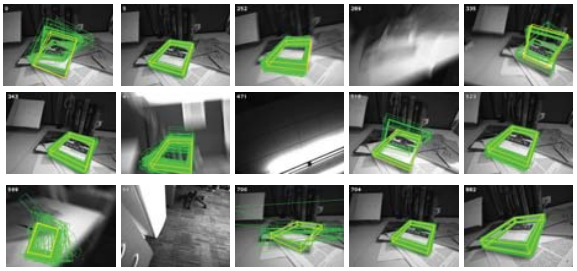
□: $N = 1$, □: $N = 100$



□: $N = 100$, □: $N = 100$, RANSAC



□: $N = 100, \lambda_a = 0.0$, RANSAC, □: $N = 100, \lambda_a = 0.5$, RANSAC



□: particles, □: mean of particles

Conclusion

- For fast particle convergence, we employed keypoint features and initialized particles by using a linear time non-iterative solution for the PnP problem.
- Particles are propagated by the state dynamics which is given by the AR process on the SE(3), and the state dynamics distributed particles more effectively.
- Edge correspondences were first enhanced by considering orientations of the edge points, and they were further refined through the RANSAC process.
- During the tracking, the proposed system appropriately re-initialized by itself when the number of effective particles is below a threshold.

References

[1] J. Kwon, F. C. Park Visual tracking via particle filtering on the affine group *IJRR*, 29(2-3):198, 2010.

Affordance Prediction via Learned Object Attributes

Tucker Hermans Advisor: Aaron Bobick Co-Advisor: James M. Rehg
Computational Perception Lab - School of Interactive Computing

Affordance Prediction

- Wish to predict affordances distally from visual input


- Affordances are action possibilities defined jointly between an agent and its environment

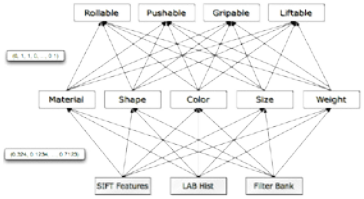
Attribute Model

- Fully supervised procedure
- Learn semantic attributes from visual features
- Learn affordances from attribute feature vectors
- Attributes share visual information across different affordances

Affordances:

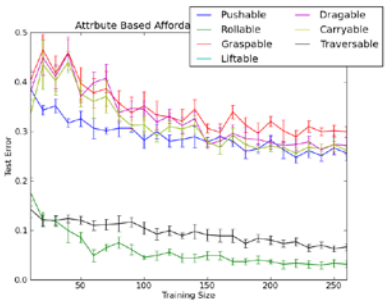
Attributes:

Visual Features:

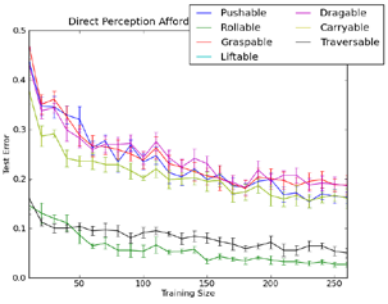


Preliminary Results

- Attribute based affordance prediction:



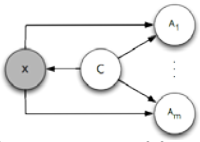
- Direct perception results:



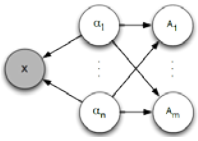
- Attributes have potential for inferring affordances of novel objects
- Novel object inference requires a rich set of objects spanning the attribute set
- Category based affordance prediction captures certain semantic biases

Future Work

CA-Full Model:

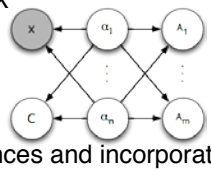


Att-Aff Model:

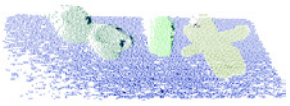


- Incorporate object category-affordance methods in a fully Bayesian framework

Combined Model:



- Learn a wider set of affordances and incorporate depth features using PR2




Robotics@GT & Intelligent Machines

Georgia Tech College of Computing

Fully Distributed Multi-Robot Simultaneous Localization and Mapping

Alex Cunningham and Frank Dellaert
Center for Robotics and Intelligent Machines, Georgia Institute of Technology, Atlanta GA, 30332

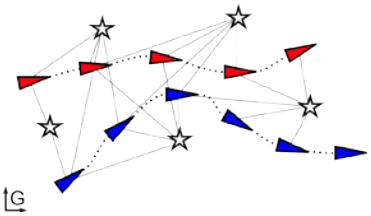


Figure 1. An example scenario with two robots driving through an environment with landmarks (stars).

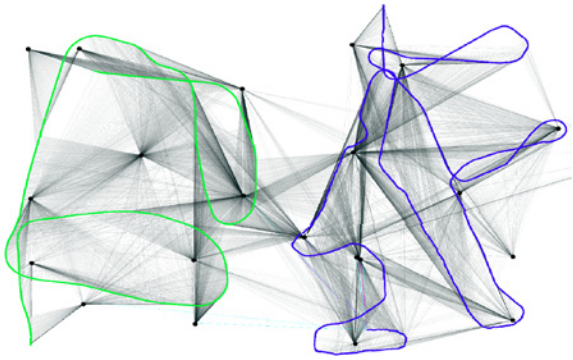


Figure 2. Two robots with globally consistent landmarks (black circles), showing corrected trajectories (green and blue lines) and landmark observations (translucent lines), shown after aligning landmarks and global optimization.

Abstract

Multi-robot SLAM systems are necessary to coordinate teams of robots by producing consistent, reliable maps of the environment. One challenge in a multi-robot system not present in single robot SLAM is finding globally consistent labels for landmarks observed by separate robots when the starting reference frames of the robots are not known. We present a novel, RANSAC-based, approach for performing the between-robot data associations and initialization of relative frames of reference, obtaining an end-to-end multi-robot SLAM system, when combined with our previous DDF-SAM approach, for which have only shown simulated result until now.

Overview

The primary requirements for multi-robot mapping system useful in harsh environments, which performs Decentralized Data Fusion (DDF), are as follows:

- Scalable in computational cost
- Scalable in communication bandwidth as the number of robots increases
- Robust to node failure
- Robust to changes in network topology

DDF-SAM

The DDF-SAM system, introduced in [1] and expanded in [2], consists of three main modules:

1) **Local Mapping Module**: Performs full nonlinear SAM to solve for the full trajectory and landmark map, then compresses the local map to broadcast to neighboring robots.
2) **Communications Module**: Updates a cache of compressed maps from many robots with correspondences and initializations and computes multi-robot data associations.
3) **Global Mapping Module**: Optimizes graph over all known neighbors and yields a global feature map.

Local Smoothing and Mapping (SAM) solves nonlinear least-squares optimization problem, while global optimization introduces *hard equality constraints* to bind landmarks in different reference frames together.

Map Matching

- Rather than matching landmarks directly, we compute features on the landmarks
- We compute *Delaunay triangulations* (shown in Fig. 3) over the landmarks in a map
- Correspondences between the triangles will be more robust to noise
- We use these triangle correspondences as inputs to a RANSAC algorithm to find data associations and relative reference frames

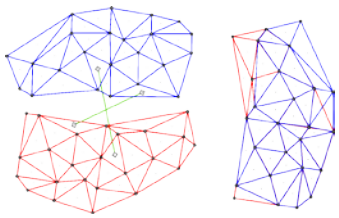


Figure 3. Map matching via triangulations of feature maps. Left: two sets of landmarks, their triangulations, and correspondences. Right: matched maps overlaid.

Experiments

We tested the system in the following scenario

- Three robots (shown in Fig. 4) equipped with laser scanners
- Parking lot environment with pole features added
- Two runs through the environment, manually controlled by human

All results were computed off-line for visualization, using the *GTSAM* graphical inference library for optimization. Full map outputs can be seen in Fig. 5, with a closeup of a two-robot case in Fig. 2.

Figure 4. Robots used for experiment and example pole feature (left) and aerial view of the parking lot used in Freiburg.

Figure 5. Full trajectories and global landmarks with 3 robots in Freiburg dataset for run 1 (top) and run 2.

Conclusions

- Decentralized multi-robot maps are feasible with real-time performance
- We can assemble a global map even when the initial positions of the robots are unknown
- The system scales with the number of robots while producing accurate maps, as in Fig. 2 and Fig. 5

Acknowledgements

Freiburg experiments and the data association approach come from joint work with Kai Wurm and Wolfram Burgard at the University of Freiburg.

References

[1] A. Cunningham, M. Paluri, and F. Dellaert, "DDF-SAM: Fully distributed SLAM using constrained factor graphs," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2010.
[2] A. Cunningham, K. Wurm, W. Burgard, and F. Dellaert, "Fully distributed smoothing and mapping with robust multi-robot data association," submitted to *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2011.

Robotics@GT & Intelligent Machines

Visual Perception for Mobile Robot Navigation

Richard Roberts, Duy-Nguyen Ta and Frank Dellaert

Goals

- Robust and reliable perception for mobile robots.
- Robot motion and environment structure.

Sting

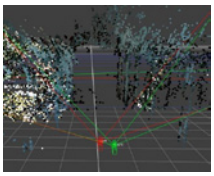
ARDrone

LAGR

- Transportation, search and rescue, planetary exploration, military, police.

Problems

- Point-cloud perception is of limited usefulness to mobile robots.



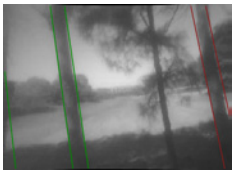
- Would like to use many cameras.
- Real-time computation required.
- Smaller robots have limited computation.

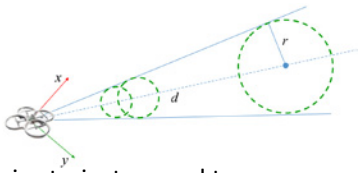
Perception for Quadrotor Navigation

Roberts, Nguyen, Straub, and Dellaert. *Under review.*

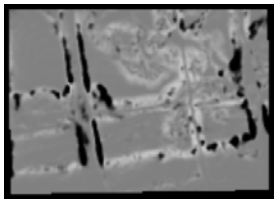
Model-based detection and tracking of trees.

Maximize image gradients on predicted tree edges

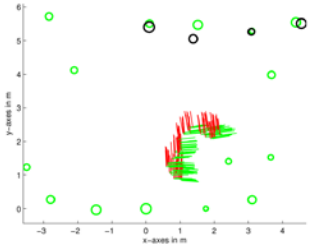


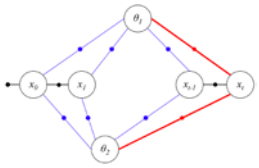


Detect only nearby trees using motion saliency



Simultaneously optimize trajectory and trees

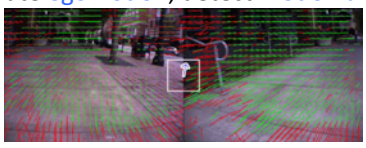




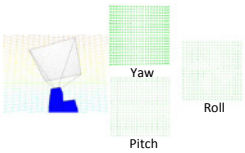
Learning General Optical Flow Subspaces

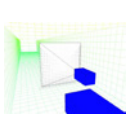
Roberts, Potthast, and Dellaert. CVPR 2009.

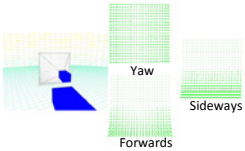
Estimate egomotion, detect motion anomalies



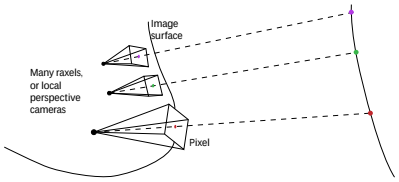
Optical flow subspaces

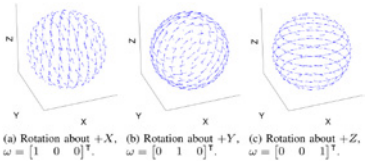






General optics





Robotics@GT & Intelligent Machines

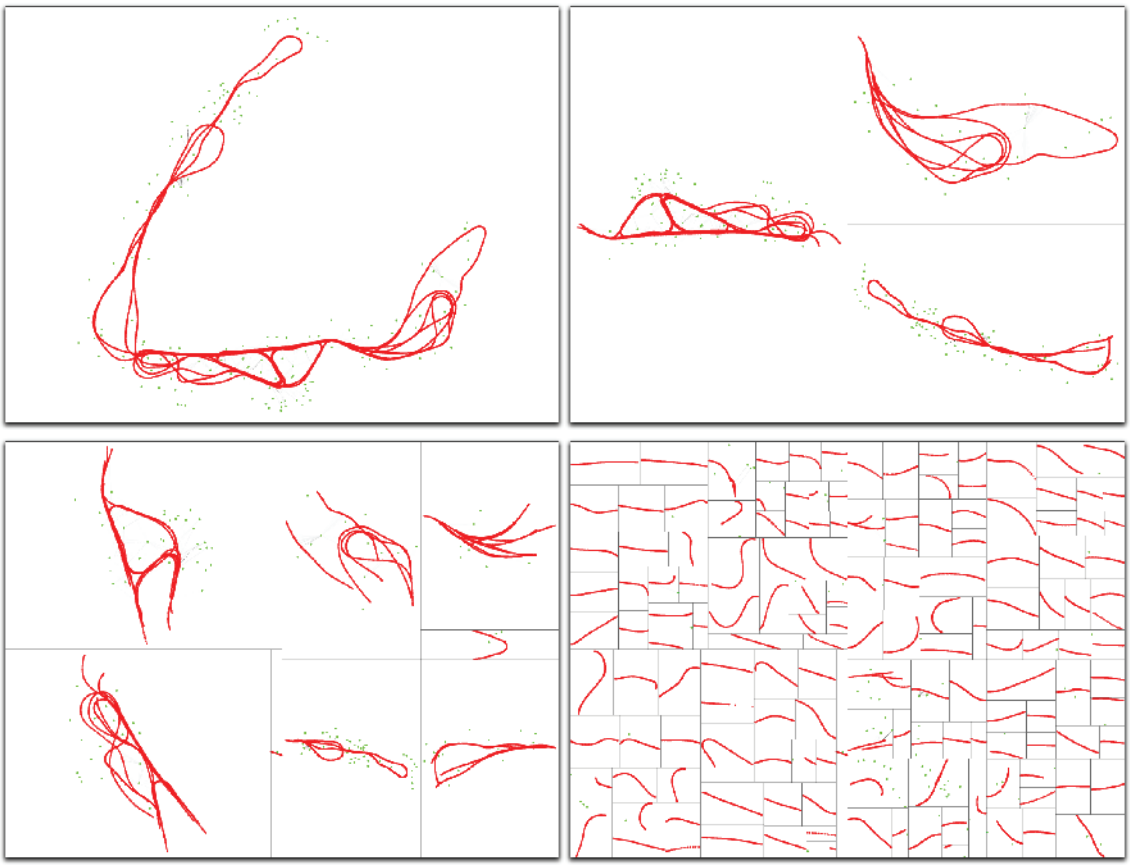
Tectonic Smoothing and Mapping

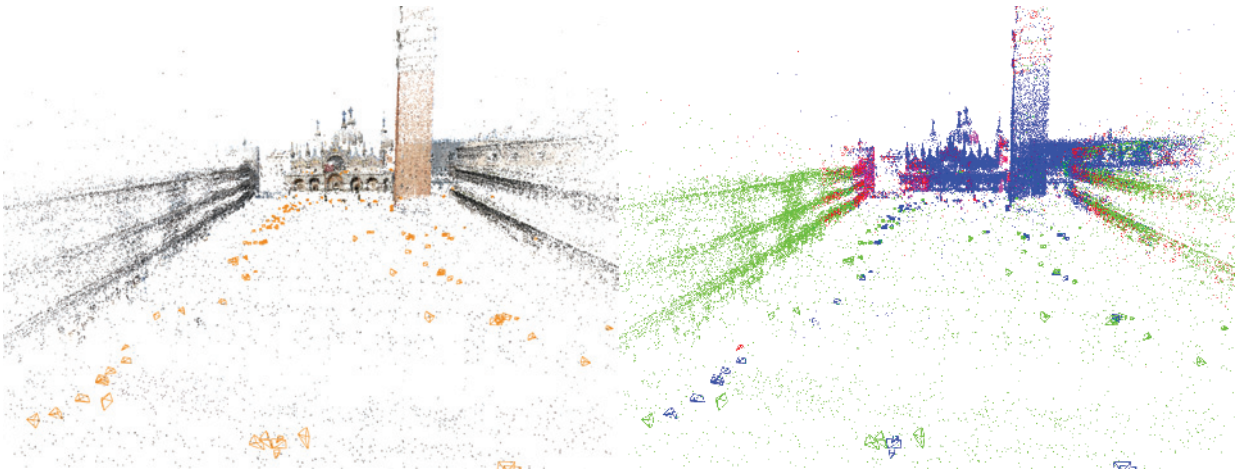
Kai Ni and Frank Dellaert

Divide-and-Conquer

Solve large-scale mapping problems in a fast, exact, robust, and scalable way.

Recursively partition the data using nested dissection algorithm and then solve from bottom up.



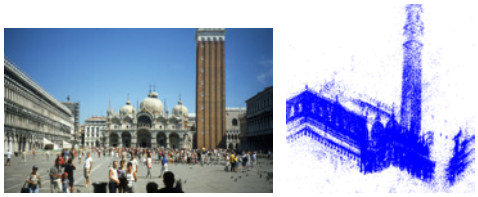


Subgraph Preconditioned Conjugate Gradient for Large-Scale Bundle Adjustment

Yong-Dian Jian, Doru C. Balcan and Frank Dellaert

Introduction

- Reconstructing 3-D structure from images is an interesting problem (SfM)



- In the SfM pipeline, we are interested in Bundle Adjustment, which serves to optimize the cameras and points at the final stage

Bundle Adjustment (BA)

- Nonlinear Optimization Problem:

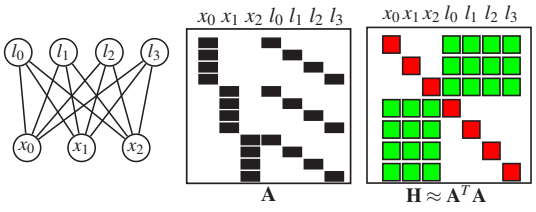
$$\min_{x,l} \sum_{k=1}^{K} \|h_k(x_{ki}, l_{kj}) - z_k\|^2 \qquad (1)$$

x: camera, l: 3-D point, z: measurement, h: projection function

- Linearize Eq. (1) to derive a linear system, and solve it for $\delta\mathbf{x}$ (the concatenation of all δx and δl)

$$\mathbf{A}\delta\mathbf{x} = \mathbf{b} \quad \text{(least squares form)}$$
$$\mathbf{A}^T\mathbf{A}\delta\mathbf{x} = \mathbf{A}^T\mathbf{b} \quad \text{(normal form)}$$
$$\delta\mathbf{x} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}$$

Graph Perspective



- Solving graphs is equivalent to solving linear systems
- Properties of the BA graphs:
  - Bipartite: No edges between cameras or points
  - Unbalanced: Much more points than the cameras

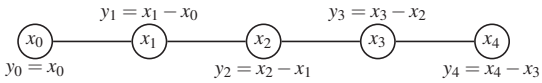
Preconditioning

- If $\mathbf{H}$ is ill-conditioned, we can solve an equivalent equation:

$$\mathbf{S}^{-T}\mathbf{H}\mathbf{S}^{-1}\delta\mathbf{y} = \mathbf{S}^{-T}\mathbf{c}$$

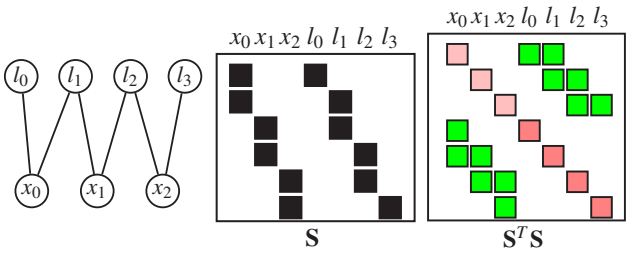
  - $\mathbf{S}^{-1}$ shoulbe easy to compute
  - $\mathbf{S}^{-T}\mathbf{H}\mathbf{S}^{-1}$ is well-conditioned

- Variable Reparametrization

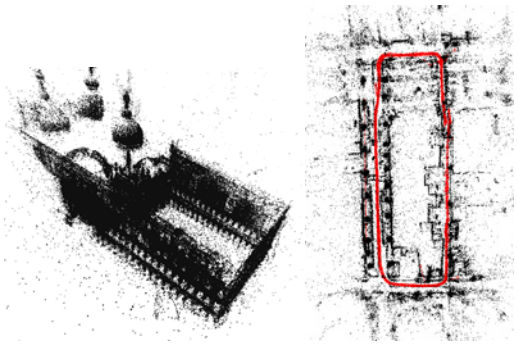


Subgraph Preconditioner

- Main Idea: Take a subgraph as the preconditioner
- An example:



Results



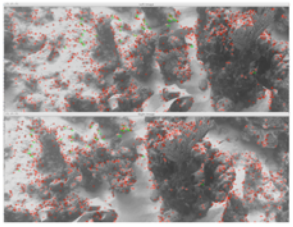
Robotics@GT & Intelligent Machines

Sparse 3D Mapping of Underwater Environments

Chris Beall and Frank Dellaert

MOTIVATION

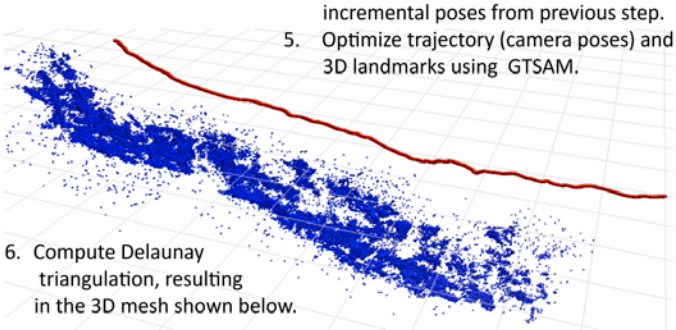
Rapid deterioration of coral reefs is a growing concern, calling for regular monitoring. Computer vision allows us to build 3D models of an environment from two-dimensional images. To construct the models, a stereo-rig is used to collect high definition videos. The construction of maps is non-trivial, and we solve this problem by employing a smoothing and mapping toolkit (GTSAM) developed in our lab.



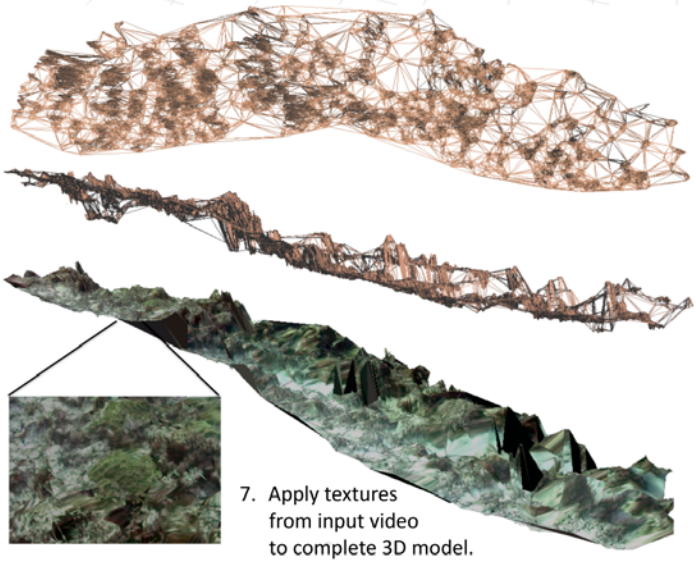
CORAL REEF, ANDROS ISLAND

METHODOLOGY

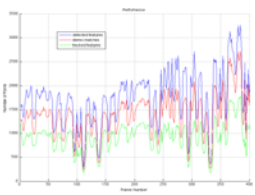
1. Rectify and synchronize stereo video.
2. Extract SURF features in each frame for stereo and temporal correspondence.
3. Use RANSAC (Random Sample Consensus) and three-point algorithms to estimate frame to frame motion.
4. Generate a trajectory (shown in red) and 3D feature point cloud using incremental poses from previous step.
5. Optimize trajectory (camera poses) and 3D landmarks using GTSAM.



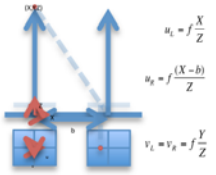
6. Compute Delaunay triangulation, resulting in the 3D mesh shown below.



7. Apply textures from input video to complete 3D model.

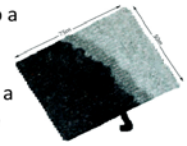


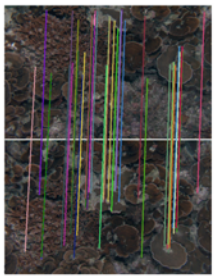
LEFT: Feature tracking performance
RIGHT: Stereo camera geometry and projection equations

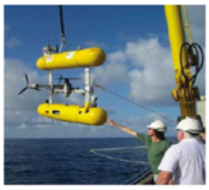


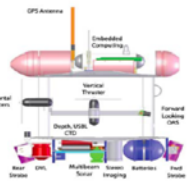
SOUTH SCOTT REEF, AUSTRALIA

We applied our smoothing and mapping algorithm to a large-scale underwater dataset collected using an autonomous underwater vehicle (AUV). 3D reconstructions have previously been obtained using a filtering approach, and in this work we show that the smoothing and mapping solution results in a more consistent map.

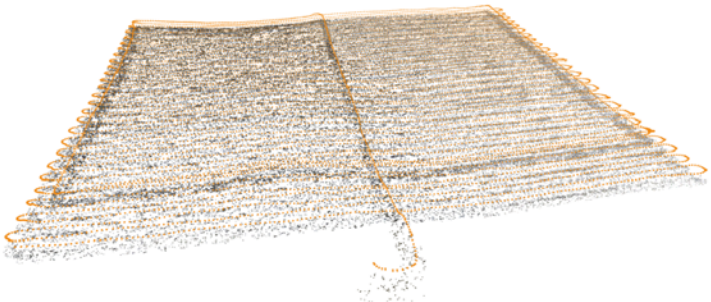




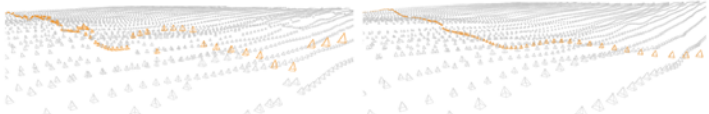




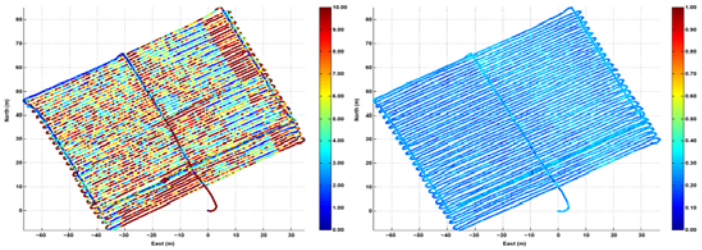
TOP: The AUV Sirius being retrieved after a mission aboard the R/V Southern Surveyor and a system diagram
LEFT: Features matched across consecutive frames of the left camera



Smoothing and mapping result consisting of 9831 camera poses (orange) and 185261 landmarks.



Partial view of AUV trajectory. The first leg north is highlighted. The smoothing and mapping result (right) is notably smoother than the filtering result (left).



Root mean square projection errors per camera, with camera locations plotted using the respective optimization results. The error is shown in color. Note the difference in error scales. Filtering result (left), and smoothing and mapping result (right).

ACKNOWLEDGEMENT
We thank Ian Mahon and Stefan B. Williams of the Australian Centre for Field Robotics for sharing their data.

REFERENCES
[1] C. Beall, B. Lawrence, V. Ila, and F. Dellaert, "3D Reconstruction of Underwater Structures," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS), 2010.*
[2] C. Beall, F. Dellaert, I. Mahon and S. Williams , "Bundle Adjustment in Large-Scale 3D Reconstructions based on Underwater Robotic Surveys," in *OCEANS 2011. Proceedings of IEEE, Santander, Spain, June 2011.*

Robotics@GT & Intelligent Machines

3D Visualization of the Operating Room Using Advanced Motion Capture: A Novel Paradigm to Expand Simulation-Based Surgical Education

Eric L. Sarin[1], Kihwan Kim[2], Irfan Essa[2], and William A.Cooper[3]

Inova Heart and Vascular Institute[1],
College of Computing, Georgia Institute of Technology [2],
Division of Cardiothoracic Surgery, Emory University School of Medicine[3]

• The Society of Thoracic Surgeons 47th Annual Meeting (STS 2011)
• International Society for Minimally Invasive Cardiac Surgery (ISMICS 2011)

Motivation

- There are multiple challenges facing the current era of cardiothoracic surgical training
- Cardiac surgical operations are trending towards older, more complex patients
- Changes to resident training programs including the mandatory 80 hour work week and the shift towards integrated programs will shorten the overall length of training
- Use of simulation techniques in cardiothoracic surgical education has been highly effective and well-received by trainees

Purpose

- To ascertain whether advanced visualization techniques can be effectively applied in the operating room
- Utilization of the video data to develop a viable simulation platform for cardiac surgical procedures.

Our Approach

Method Overview We deploy multiple cameras in the operating room. We typically utilize four fixed camera views to provide a global overview of the operating room (Fig. 1). These views are supplemented with additional views from the others cameras focused on the surgery to provide a more close-up view of the procedure itself

Figure 1

Registration Each camera view is then registered onto a reference plane in a virtual 3-dimensional space. Projective homography matrices are calculated from each view based on the reference points (Fig. 2)

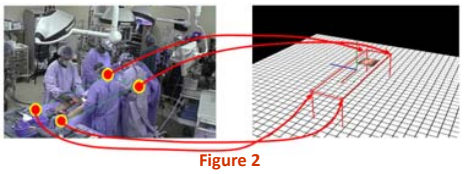


Figure 2

Segmentation A region of each person in the videos is extracted using a segmentation algorithm over time (Fig.3a,b) and the location and height of each person is measured from the videos (Fig. 3c). This information is used for locating the billboard of each individual person consistently over time in 3D space (Fig. 4). The geometric constraints of vertical vanishing points from each camera are used for this step

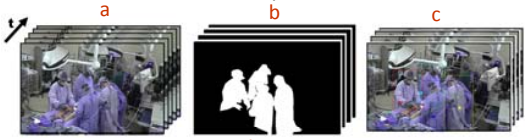


Figure 3

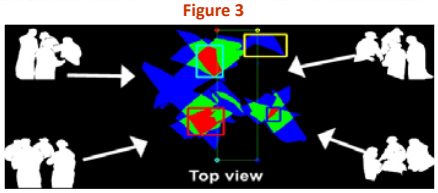


Figure 4

Video Textured Billboard The original image in each segmented region (Fig.5a,b) is directly mapped onto a billboard for each video frame. A video-textured billboard is now rendered for the virtual view point. This maintains consistent visuals for the users despite slight changes in the virtual viewpoint. (Fig.5 c). For the smooth transition of virtual view, we applied novel blending techniques with adequate depth tests to our

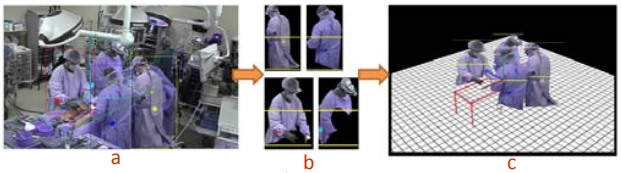


Figure 5

Results

Manipulation of the virtual viewpoint allows access to different camera inputs. Rotation from right to left (Fig. 6) brings the local view of the heart into view. The ability to rotate and magnify the view (Fig. 7) allows visualization of the surgeon's hands and technical details of the operation. Application also allows users to select a person to be (1) transparent to look through (Fig.8b), or (2) highlighted to concentrate on actions. (Fig.8c)

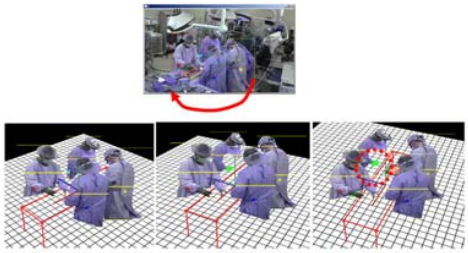


Figure 6

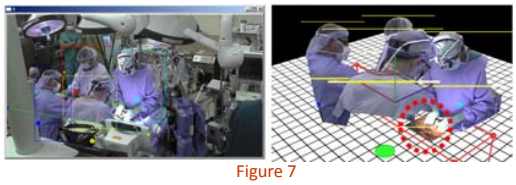


Figure 7

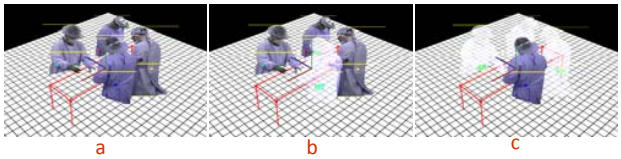


Figure 8

Conclusion

- Advanced motion capture using video based rendering can be used to provide an accurate representation of cardiac surgical procedures

- Video-based simulation represents a potentially viable adjunct to the currently available techniques utilized in cardiothoracic surgical education.

Using the Microsoft Kinect in order to Analyze Medical Exercises

Daniel Castro, Henry Dooley, Gaurav Mathur, and Irfan Essa. Computational Perception Laboratory, Georgia Institute of Technology

Abstract: Abstract: Microsoft's Kinect hardware has brought depth sensing hardware to the everyday user at an affordable price. It has set records for being one of the highest selling consumer electronics on the market today. Because this hardware is becoming increasingly accessible, we are developing new techniques to expand upon existing applications to leverage the Kinect as a medical tool. By using the Kinect to track a user's skeleton in real time, we are able to analyze how their body performs certain exercises, in order to aid in early recognition and diagnosis of physical disabilities or health problems.

After establishing a framework capable of capturing critical skeletal data, we can then apply a supervised learning algorithm to empirically classify an individual, and suggest possible causes of negative performance. This proof of concept demonstrates the collection of skeletal data, which would then be ready to be classified under the guidance of medical experts.

Diagram 1 - Original Exercise

Diagram 1: The original exercise video and the chosen frames for our four key states.

Diagram 2 - Standing up Exercise Data

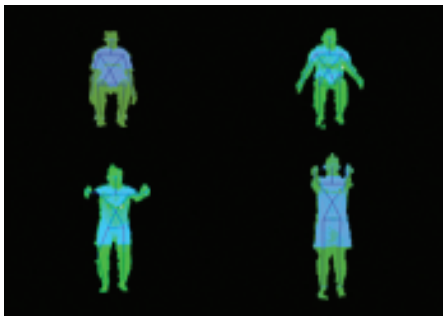


Diagram 2:The four poses that comprise the No-Bands Sit-Stand Squat Exercise.

Our Method: We begin by feeding our sensor data into the OpenNI (Natural Interactions) library, which will allow us to track our user. We then capture and analyze the angles of relevant skeletal joints. By matching the angles of the users joints against the angles of a known state, we are able to identify when a user has entered a specific stance or pose. In order to track the user's progress during the exercise, the exercise is broken down into a set of key states, which the user must move through successfully to complete the exercise.

For testing, we retrieved an exercise from arthritistoday.org, and broke it down into what we deemed were the four key states. These can be seen in diagram 1. Diagram 2 illustrates how our program recoginizes the user in the four different states. Note that the user is shaded green, indicating that he is performing the pose correctly. The requirements for our test exercise, the No Band sit-stand squat, can be seen in diagram 3.

Diagram 3 - State and Requirement Setup

Special Thanks to the Computational Perception Laboratory under the supervision of professor Irfan Essa, at the Georgia Institute of Technology.

Robotics@GT & Intelligent Machines

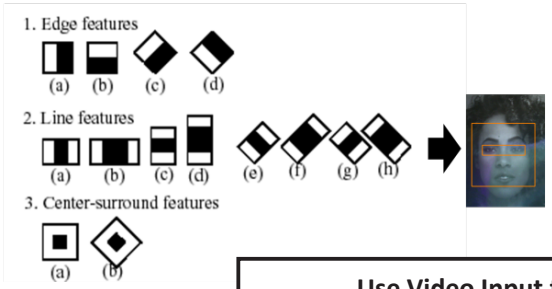
Peek-A-Boo Simon

Presenter: Joy Buolamwini
Advisers: Dr. Irfan Essa , Dr. Andrea Thomaz

Motivation: As robots begin to enter social spaces such as care centers and hospitals, there is a growing need to study Human Robotic Interaction. The objective of the Peek-A-Boo Simon initiative is to explore how to engage humans in turn taking interactions with a robotic agent to better understand how to structure these social interactions so that robots can engage meaningfully with humans in a context appropriate manner.

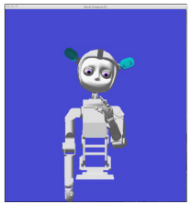
Face Detection

Video Feed → Haar-Like Features → Output



Notify Robot of Events

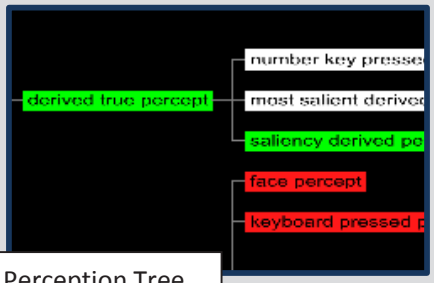
Send IRCP Packet



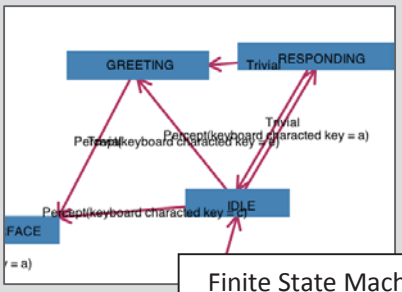
IRCP: Intra Robot Communication Protocol

Use Video Input to Construct Beliefs

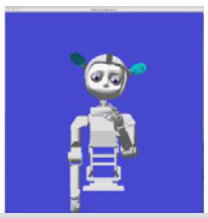
Use Beliefs to Determine Appropriate Interactions



Perception Tree



Finite State Machine

Greeting	Conceal Face	Reveal Face
		

Interactions

Motion Fields to Predict Play Evolution in Dynamic Sport Scenes

Kihwan Kim[1], Matthias Grundmann[1], Ariel Shamir[2], Iain Matthews[3], Jessica Hodgins[3] and Irfan Essa[1]

Georgia Institute of Technology[1], Interdisciplinary Center[2] and Disney Research Pittsburgh[3]

CVPR June 13-18, 2010 San Francisco, CA
IEEE CVPR 2010

Goal and Motivation

To predict how a play in a sports game will evolve, given current positions and motions of the players.

(1) Player intent can be analyzed from ground motion fields
(2) The dense ground motion field can be interpolated from a sparse set of motion vectors over time.

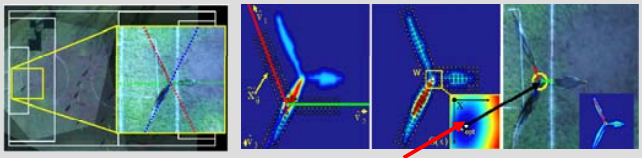
Overview of Approach

(1) Extracting players' locations on the ground from multiple cameras
(2) Generating dense flow fields from a sparse set of player motions over time using Radial Basis Function (RBF) Interpolation
(3) Detecting Points of Convergence (POC) by clustering propagated flow magnitude.

Predict the location of future play

Motion on the Ground

1. Extract ground position of each player on the field by optimizing geometric constraints from vertical vanishing points



$$G(x)=\sum_{k=0}^{N}\sum_{i=0}^{n} PC(\tilde{x}_i,k)\cdot d(\tilde{x}_{i,k},(\hat{v}_k-x)) \quad \Rightarrow \quad x_{opt}=\arg\min_{x\in W_{init}} G(x)$$

Foreground — Distances from vertical vanishing lines

2. Motion on the ground

$$x_{opt}^{t-a}=\arg\min_{x(x,y)\in W_{opt}}(G(x)^{t-a}+\beta\cdot C(x)^{t-a})$$

Color proximity

$$[u,v]^T=\frac{\partial x}{\partial t}\cong(x_{opt}^{t}-x_{opt}^{t-a})/a$$

Global Flow Field

Spatio-temporal interpolation of player flow using RBF

$$[u,v]^T=\left[\frac{\partial x}{\partial t},\frac{\partial y}{\partial t}\right] \xrightarrow{\text{Over time}} U\in\Re^{n\times 1}, V\in\Re^{n\times 1}$$

Given a collection of *n* scattered 2D points (x, y) with associated scalar velocity values (U,V), construct a smooth velocity field that matches each of these velocities at the given locations

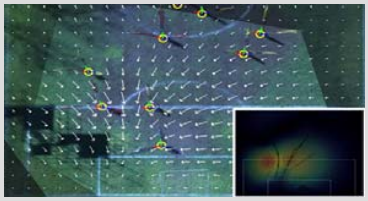
Velocity function for x direction $f(x)=c_x+\sum_{i=0}^{n}\lambda_i^x\phi(\|x-x_i\|)$

Velocity function for y direction $g(x)=c_y+\sum_{i=0}^{n}\lambda_i^y\phi(\|x-x_i\|)$

$\phi(r)=r^2\log r$ minimizes energy, $E=\int_{\Re}\left(\frac{\partial^2 f}{\partial x^2}\right)^2+2\left(\frac{\partial^2 f}{\partial x\partial y}\right)^2+\left(\frac{\partial^2 f}{\partial y^2}\right)^2 dxdy$

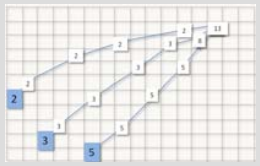
Motion field

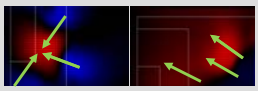
$$\Phi(x)=f(x)\mathbf{i}+g(x)\mathbf{j}=u\mathbf{i}+v\mathbf{j}$$



Points of Convergence (POC)

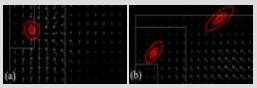
1. Given a location x(i,j), propagate the magnitude of ground flow velocities along the flow field. $\rho_{ij}^2=u_{ij}^2+v_{ij}^2$ are accumulated in the confidence table $\Psi(i+u_{ij},j+v_{ij})=\Psi(i+u_{ij},j+v_{ij})+\rho_{ij}$
2. Then, find the largest accumulation using meanshift and GMM





div = $\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}$

Red : negative divergence, Blue : positive divergence



Detected POCs
(a) Critical Points (b) Non-singular region

Results & Evaluation

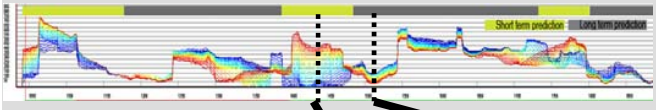
Qualitative results : Red contours : POC, White arrows : Motion Fields

Back-door play and through pass

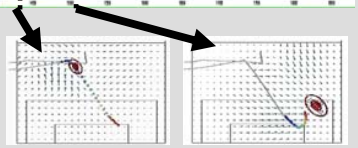
Interception and goal keeping

Faint motion and drive-in (Basket ball)

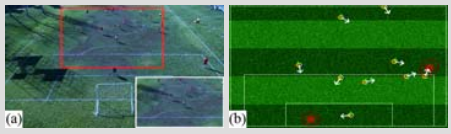
Evaluation : Distances between the location of the detected POC in the current frame vs. the location of the ball in future frames (blue : +4 frames, and red : +120 frames)



4.5 meter errors without the period that goal-keeper occupies the ball, and 6.5 meter errors for whole time offsets when we take the minima of each frame.



Applications



(a) Camera control for automated broadcasting
(b) Novel type of sports scene visualization

Conclusion

We introduce a novel approach for play evolution analysis, and show that global analysis of local players' movement can be used for predicting future events in dynamic sport scenes.

Future work : (a) Developing more efficient ways to calculate motion fields (b) Automatically driving robotically controlled cameras using our algorithm

Crowd-sourced Video Annotation by Leveraging Server-based Segmentation

Kevin Hampton, Matthias Grundmann, Irfan Essa
School of Interactive Computing

Research Goal

- Extract semantic, spatio-temporal regions representing individual objects

Possible Applications

- Ability to train semantic segmentation
- Database of annotated videos for future research

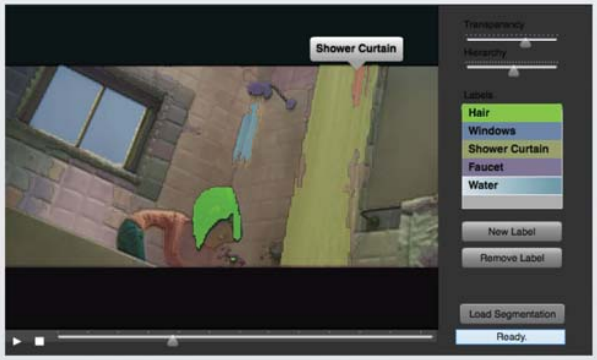
Challenges

- Creation of web interface
- Annotation of spatio-temporal regions
- Repetition of selecting regions

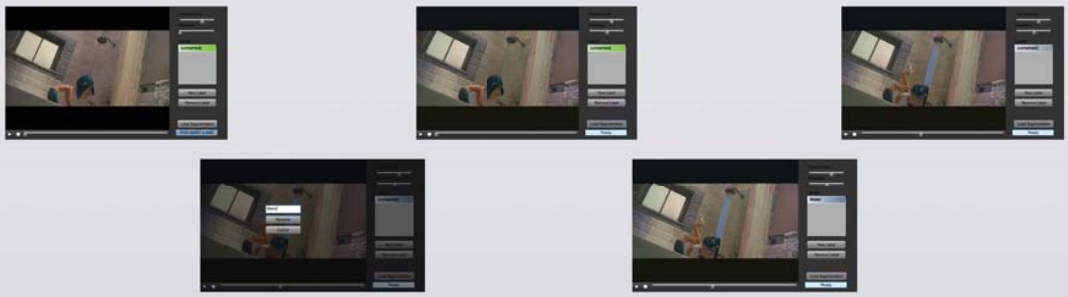
Resources

- Efficient Hierarchal Graph-based Video Segmentation (CVPR 2010 Grundmann, Kwatra, Han, Essa)
- Adobe Flash Professional
- Google Protocol Buffer

Results



Shown above is the result of an annotated video using our solution.



Shown above is the labeling process. First the user must load the segmentation.
The segmentation is then displayed. Then the user can select regions and

Approach

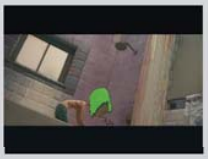
1. Creation of Flash user interface

Simple video display abilities
Integration of Precomputed Segmentation

2. Creation of user interface actions to annotate video

Selection and annotation of regions with hierarchy and transparency options

3. Trimming functionality



Region trimming functionality of spatio-temporal regions

4. Future work

Amazon Mechanical Turk
Annotation of 500 - 1000 videos
Each video annotated by 10 users

Augmenting Aerial Earth Maps with Dynamic Information

Kihwan Kim, Sangmin Oh, Jeonggyu Lee and Irfan Essa
College of Computing, Georgia Institute of Technology
{kihwan23, sangmin, glaze, irfan}@cc.gatech.edu

Oral presentation at ISMAR 2009
Virtual Reality Journal Springer 2011

Goal and Motivation

To make Augmented Aerial Earth Maps that can visualize an alive and dynamic scene within city using distributed videos

(1) Current Virtual Earth updates spatial information over 4~6 months : almost static

(2) Making use of videos can be a solution, but videos are sparsely located, and have narrow Field of View (FOV)
: Need to find spatio-temporal correlations between videos taken at different locations

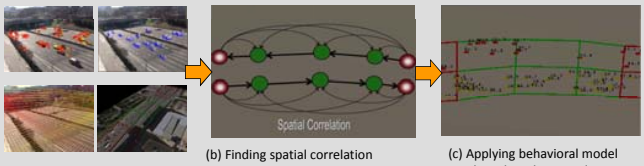
Our Approach

Four different scenarios are proposed, based on sparsity of cameras and motion complexity
: Different spatial correlations are used for visualization of the arbitrary viewing-angle and unobserved region.

1. Pedestrian visualization : single view and simple motion

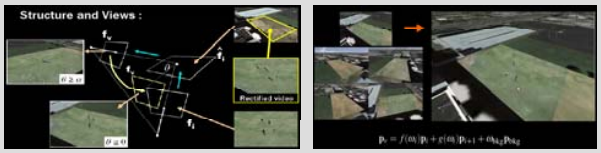
(a) Tracking pedestrian (b) Matching motion capture data (c) Projecting models onto virtual map

2. Traffic visualization : sparse views and simple motion



(a) Tracking cars and registering a geometry of the road (b) Finding spatial correlation between observed region (red) and unobserved region (green) (c) Applying behavioral model based on the spatial correlation between nodes

3. Sport visualization : overlapping views and complex motion



(a) Each rectified video is mapped onto the virtual plane, then searching two closest views based on a viewing angle of virtual view (b) A video of the arbitrary view can be reconstructed by global blending of closest views and their backgrounds

4. Sky visualization : sparse views and complex motion

(a) From sky videos taken at different locations, cloud density at in-between region can be estimated (b) Variational approach using Radial Basis Function is used for interpolating gradients of textures from observed videos (a) Propagated gradients are used to generate clouds procedurally

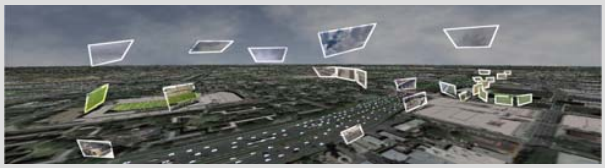
Conclusion

(1) A novel framework for visualizing dynamic scenes is proposed : Global view-blending, Parameterized behavioral model and Vision-driven procedural rendering of clouds

(2) The prototype system handles maximum 32 observations at a time at 10 frames per seconds

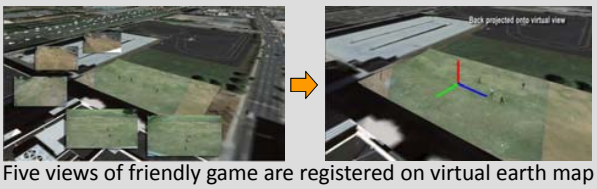
(3) The system can be used for : (a) Dynamic online map services (b) Interactive sports broadcasting (c) Surveillance applications

Results

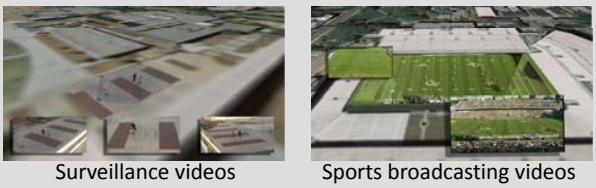


Overview of the prototype system (28 videos are used)

1. Visualization of overlapping multiple videos

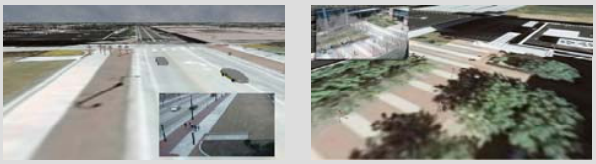


Five views of friendly game are registered on virtual earth map



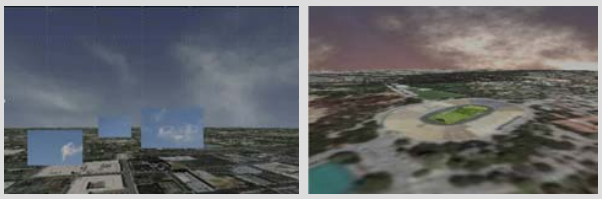
Surveillance videos Sports broadcasting videos

2. Pedestrian Visualization using single views



3. Traffic Visualization using sparse multiple videos

4. Sky/Cloud Visualization using sparse multiple videos



Sky and clouds are visualized from sparse set of sky videos Sports game can be mapped in different location from user interface

Future work : (a) aim to overcome algorithmic trade-offs (b) plan to apply other situations (i.e. river, seashore, etc.)

- For more information :
http://www.cc.gatech.edu/cpl/projects/augearth

IEEE CVPR 2010,
San Francisco

Google research

Efficient Graph-Based Hierarchical Video Segmentation

Matthias Grundmann[1,2], Vivek Kwatra[2], Mei Han[2], Irfan Essa[1]
[1]School of Interactive Computing - Georgia Institute of Technology [2]Google Inc.

Georgia Tech College of Computing
GVU Georgia Tech Robotics@GT Intelligent Machines

Research Goal

- Solve fundamental vision problem of segmentation for *unstructured* video robustly and efficiently
- Compute *spatio-temporal* regions grouped by appearance and motion similarity

Possible Applications

- Video content analysis
- Rapid, temporal rotoscoping
- Video Tooning

Challenges

- Temporal coherence
- Initial outline unavailable
- Scalability / Memory constraints

Image segmentation applied to each frame independently produces unstable segmentation results, since even small frame-to-frame changes cannot be expressed as a continuous function.

Contributions

- Volumetric approach with flow-based temporal neighborhood
- Scalable algorithm using out-of-core and clip based processing
- Hierarchal segmentation scheme, choose granularity *after* segmentation

Results

Spatio-Temporal video segmentation for various test cases.

Shown are two frames from a video sequence and our obtained segmentation. Region identity is indicated by color. Segmentation exhibits consistent region identity and boundaries under conditions such as significant motion (water skier doing 720° turn, first row), illumination changes (explosion in the background, middle row) and camera motion (panning down, last row). Video length varies from 2 - 6 seconds.

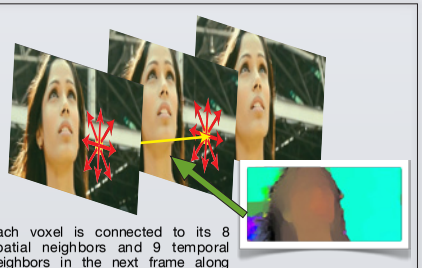
Applications - User selected regions and video tooning

Segmentation applied to long sequences (~ 30 seconds)

Results for clip based processing. By splitting the video into overlapping clips of 30 frames and constraining adjacent solutions to be temporally coherent, we can segment consecutive video clips beyond 40 seconds. Top row: Segmentation for bar-scene from Goodfellas. Bottom row: Yu-Na Kim, 2009 World Championships. Ice skater (green) selected by only a single click, Olympus sign (pink) selected by two.

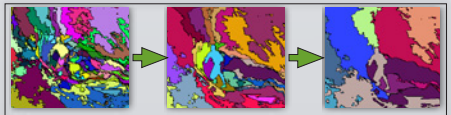
Approach

1. Construct volumetric graph and apply graph-based segmentation algorithm from [1].



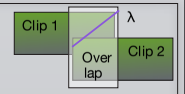
Each voxel is connected to its 8 spatial neighbors and 9 temporal neighbors in the next frame along dense optical flow (we used GPU dense flow from [2]). Edge cost is defined as difference in gamma corrected RGB space.

2. Apply hierarchical segmentation → tree of super-regions



The initial volumetric segmentation results in a spatio-temporal over-segmentation (top-left), which is used to construct a so called region graph (right inset). The region graph represents each region by a node, connected to its neighbors by edges. Edge cost is derived from local region descriptors capturing appearance and motion. Applying graph based segmentation iteratively, we obtain a hierarchy of super-regions (top middle, right).

3. Clip based processing with graph constraints



We split the video into clips of 30 frames with an overlap of 10 frames. Here, the segmentation of clip 1 is used to re-weight the edge weights in the overlap region (times λ for edges within regions, times λK across boundaries, with λ varying linearly from 0 to 1 across the overlap, K being a large constant). Consequently the segmentation of clip 1 and 2 agree on the first frame in the overlap but vary on later frames effectively giving the segmentation of each clip some independence → robustness.

References

[1] **Efficient Graph-Based Image Segmentation**
Pedro F. Felzenszwalb and Daniel P. Huttenlocher
International Journal of Computer Vision, September 2004
[2] **Anisotropic Huber-L1 optical flow**
M. Werlberger, W. Trobin, T. Pock, A. Wedel, D. Cremers and H. Bischof
BMVC, London, UK, 2009.

IEEE CVPR 2010,
San Francisco

Google research

Discontinuous Seam-Carving for Video Retargeting

Matthias Grundmann[1,2], Vivek Kwatra[2], Mei Han[2], Irfan Essa[1]
[1]School of Interactive Computing - Georgia Institute of Technology [2] Google Inc.

Georgia Tech College of Computing
GVU Robotics@GT

Research Goal

- Develop efficient algorithm for content-aware, *real-time* resizing of *streaming* video

Possible Applications

- 4:3 <-> 16:9 conversion
- object removal
- full-screen streaming to mobile clients (various resolutions)

Challenges

- Enforcing temporal coherence
- Adapting to fast camera motions and sudden occlusions
- Preserve shape of objects during iterative seam-carving process
- Allow for rapid user interaction to define important details

Contributions

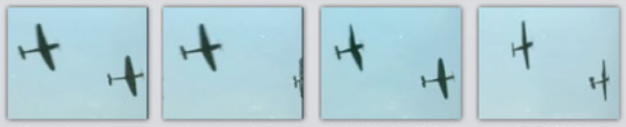
- Optimize temporal coherence in the retargeted video in contrast to imposing geometrical constraints, e.g.surfaces.
- Shape-preserving *gradient-variation* based spatial coherence
- Optional spatio-temporal saliency:
  a) Averaging frame-based saliency over regions
  b) User specified region brushing

Results

Retargeting Results for YouTube-Video "Apologize".

Shown are six frames from the result of our retargeting algorithm applied to a sub-clip of the video "Apologize", © 2006 One Republic. Original frame is shown on the left, retargeted result on the right. Shot boundary processing is used beforehand to separate the individual shots before processing.

Shape preserving optimization



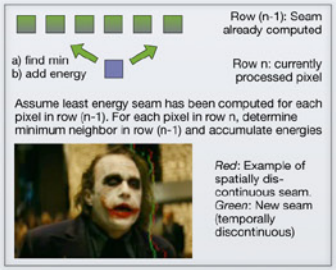
Effect of our spatial coherence measure shown for one still frame from Valkyrie, © 2007 MGM

Support for user-supplied saliency

Camera motion and scene complexity requires saliency definitions beyond what may be automatically computable. In this example, removing the white pillar introduces significant motion artifacts (camera zooms out).

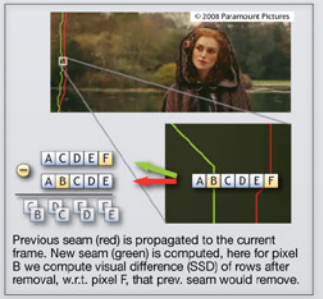
Approach

- Resize operator: Seam removal and duplication (see [3])
- Iterative computation of *discontinuous* minimal-energy seams via dynamic programming
- Energy is weighted sum of:
  a) Saliency: Importance measure
  b) Temporal coherence
  c) Spatial coherence: Shape preserving



Temporal coherence

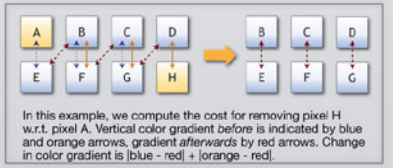
- General idea: Instead of the retargeting operators, retargeted frames should be visually similar
- Propagate previous seam to next frame (most temporally coherent)
- Compute new seam such that visual difference (SSD) of result to most temporally coherent one is minimized



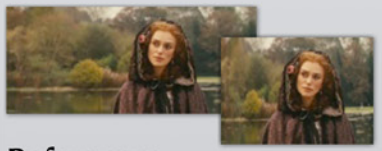
Previous seam (red) is propagated to the current frame. New seam (green) is computed, here for pixel B we compute visual difference (SSD) of rows after removal, w.r.t. pixel F, that prev. seam would remove.

Spatial coherence

- Measure change in color gradient
- Vertical gradients become diagonal



In this example, we compute the cost for removing pixel H w.r.t. pixel A. Vertical color gradient *before* is indicated by blue and orange arrows, gradient *afterwards* by red arrows. Change in color gradient is |blue - red| + |orange - red|.



References

[1] **Improved Seam Carving for Video Retargeting**
Michael Rubinstein, Ariel Shamir and Shai Avidan, SIGGRAPH, 2008

[2] **Non-homogenous Content-Driven Video Retargeting**
Lior Wolf, Moshe Guttmann and Daniel Cohen-Or, ICCV, 2007

[3] **Seam Carving for Content-Aware Image Resizing**
Ariel Shamir and Shai Avidan, SIGGRAPH, 2007

IEEE CVPR 2011,
Colorado Springs

Google research

Auto-Directed Video Stabilization with Robust L1 Optimal Camera Paths

Matthias Grundmann[1,2], Vivek Kwatra[2], Irfan Essa[1]
[1]School of Interactive Computing - Georgia Institute of Technology [2]Google Inc.

Georgia Tech College of Computing
GVU Robotics@GT & Intelligent Machines

Research Goal

- Stabilization of casual web videos w.r.t. cinematographic principles

Requirements / Properties

- Fully automatic / no user interaction
- Robustness (foreground motion, video quality, sudden scene changes)
- No failures (fallback to shake original)
- Faster than real-time on low-res video

Post-Process Video Stabilization

- 3 main steps:
  - Estimate camera motion (tracking)
  - Smooth camera path
  - Synthesize from new viewpoint

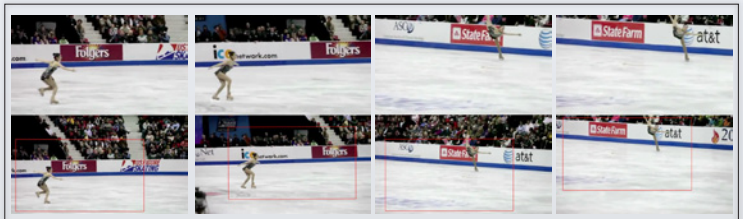
Contributions

- Novel technique for L1 optimal camera path smoothing
  - Unified framework for Video Stabilization and Video Retargeting
  - Directed by variety of constraints
- State-of-the-art results while
  - Computational cheaper
  - Larger applicability
- Fall-back strategy to shaky original

Video Retargeting Application

- Assume source video is stable ($F_t = I$)
- Steer crop window by motion saliency

Stabilization Results

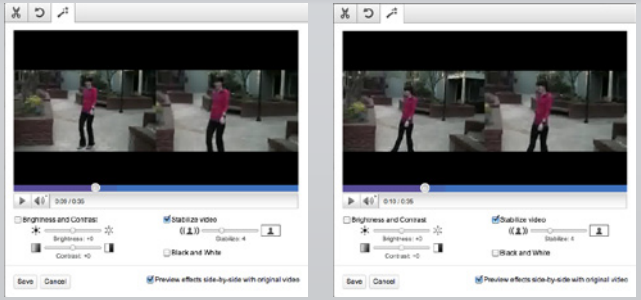


Stabilization (top) for casually shot web video (bottom). We perform stabilization by computing an optimal crop window (red) w.r.t. a cinematography inspired objective and subject to several constraints (e.g. crop window has to stay within frame at all times).

Auto-Directed Video Stabilization driven by saliency constraints using a face detector. Top: Face-directed result, optimized to stabilize video and center face. Middle: Stabilization without face constraints. Bottom: Original with crop window overlaid.
Notice, how the up and down motion of the camera is completely eliminated in both results. In addition, the saliency-driven results centers the object of interest (jumping girl) in the middle of the crop window.

Live on YouTube (real-time)

youtube.com/editor



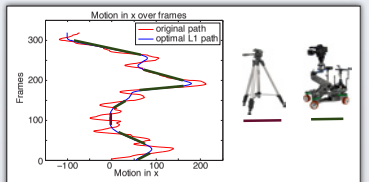
Real-time implementation in YouTube Video Editor. After a short delay (< 3 s) the stabilized result is streamed in real-time to the user making it easy and interactive to select appropriate stabilization strength (crop size). Results can be published, i.e. re-rendered in full HD (up to 1080p).

Approach

- Optimal partition (via L1) of stable path P(t) into segments
  - Tripod: $DP(t) = 0$
  - Dolly: $D^2P(t) = 0$
  - Transition: $D^3P(t) = 0$
- L1 minimization:

$$w_1 |DP|_1 + w_2 |D^2P|_1 + w_3 |D^3P|_1 \quad \text{subject to constraints}$$

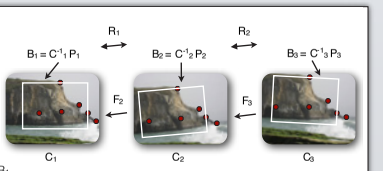
- L1 to LP: $\min |Ax|_1 \Rightarrow \min e \geq 0 \quad \text{s.t.} \quad -e \leq Ax \leq e$



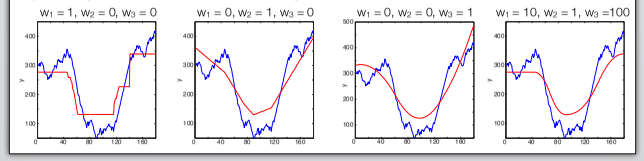
Optimal path (blue) fitted to original path (red). Optimal path is only composed of static (tripod model), linear (dolly shot or pan) and parabolic (ease in and out transitions) segments, corresponding to first, second and third derivative being zero.

Derivation for DP

- Known frame transform $F_2 \Rightarrow C_2 = C_1 F_2$
- Optimal path P_t:
  - $DP_2 = P_2 - P_1$
  - $P_1 = C_1 B_1$ original C_1 + update B_1
  - $DP_2 = C_2 B_2 - C_1 B_1$
  - Minimize L1 of residual $DP_2 = C_1F_2B_2 - C_1B_1 \rightarrow R_1 = F_2B_2 - B_1$



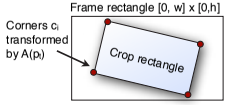
Optimal paths (red) for various weight combinations



Constraints

- Inclusion [crop window has to stay within frame bounds, hard, two-sided]
- Proximity constraint [preserve intent, bounds on scale & rotation, hard]
- Saliency constraints [crop window has to cover salient points s_i, soft]

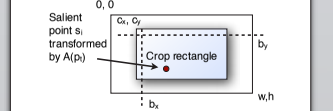
Inclusion constraint



Inclusion constraint for affine transformation $A(p_t)$:

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} \leq \underbrace{\begin{pmatrix} 1 & 0 & c_i^x & c_i^y & 0 & 0 \\ 0 & 1 & 0 & 0 & c_i^x & c_i^y \end{pmatrix}}_{:=CR_i} p_t \leq \begin{pmatrix} w \\ h \end{pmatrix}$$

Saliency constraint



Similar to inclusion constraint, but one sided soft constraint, minimize slack ϵ

$$\begin{pmatrix} 1 & 0 & s_i^x & s_i^y & 0 & 0 \\ 0 & 1 & 0 & 0 & s_i^x & s_i^y \end{pmatrix} p_t - \begin{pmatrix} b_x \\ b_y \end{pmatrix} \geq \begin{pmatrix} -\epsilon_x \\ -\epsilon_y \end{pmatrix}$$

Robotics@GT & Intelligent Machines

Understanding Egocentric Activities

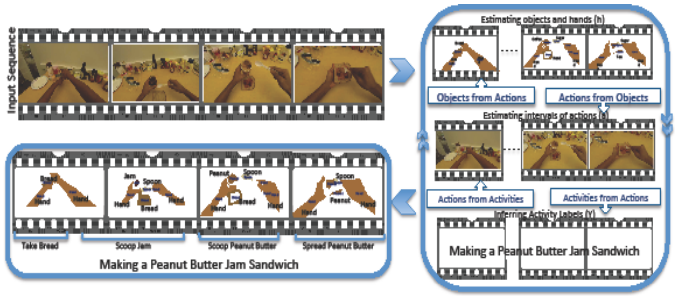
Georgia Tech College of Computing School of Interactive Computing

Alireza Fathi, Xiaofeng Ren, Ali Farhadi, James M. Rehg

School of Interactive Computing , College of Computing, Georgia Institute of Technology

{afathi3,rehg}@gatech.edu
xiaofeng.ren@intel.com
afarhad2@uiuc.edu

Motivation



Goal

• Leverage the structure of the egocentric domain to segment the image into spatial (hands, object, background) and temporal (actions) parts.

Contributions

- Introduce a weakly supervised method for object recognition.
- Introduction of a novel semantic feature representation for actions based on object and hand interactions.
- A joint model of objects, actions and activities.

Object Recognition

frame | left hand | right hand | object

$$Pr(x_i|B_j) = s(x_i, B_j) = max_{x_k \in B_j} exp\left(-\frac{\| x_i - x_k \|^2}{\sigma^2}\right)$$

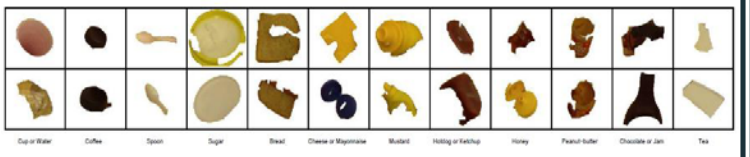
$$\begin{bmatrix} s(x_1,B_1) & \cdots & s(x_1,B_t) \\ s(x_2,B_1) & & s(x_2,B_t) \\ \vdots & & \vdots \\ s(x_m,B_1) & \cdots & s(x_m,B_t) \end{bmatrix}$$

$$min_{w,b,\xi,\eta}\left\{\sum_{c=1}^{r} \mid w_c \mid + C\sum_{j=1}^{l}\xi_j\right\} \quad (2)$$

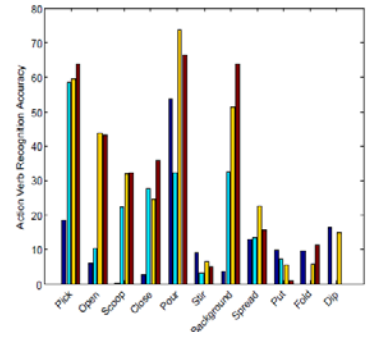
$$w_{y_i}.s(:,j) \geq w_c.s(:,j) + \delta_{y_i,c,B_j} - \xi_j \; s.t. \; \forall c = \{1,...,r\}$$

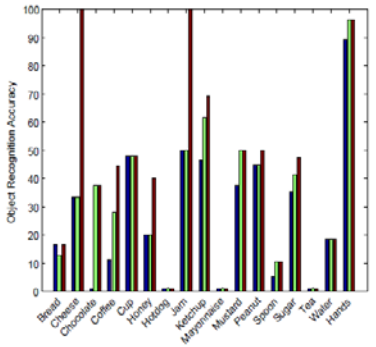
$$w_c(p) = w_c(q) \; s.t. \; \forall c = \{1,...,r\}, \; (p,q) \in \mathcal{C}$$

$$\xi_j \geq 0 \; s.t. \; \forall j = \{1,...,l\}$$



Results





Robotics@GT & Intelligent Machines

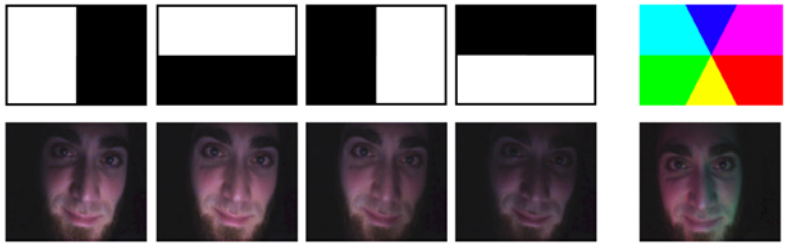
Photometric Stereo via Computer Screen Lighting for Real-time Surface Reconstruction

Grant Schindler
Georgia Institute of Technology

Problem

Using only the light from a computer screen to illuminate an object from multiple directions, we use photometric stereo techniques to recover a textured 3D surface in real-time from live video. This technique is viable on any laptop or desktop computer with a screen and webcam.

Computer Screen Lighting

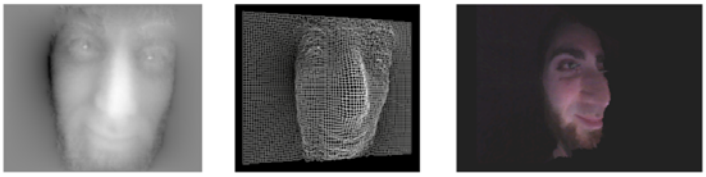


Recovering Surface Normals

Z X Y

Singular Value Decomposition. Images live in a 3-dimensional subspace defined by dominant eigenvectors corresponding to the z, x, and y components of the surface normal at every pixel in the image.

Surface Reconstruction



Depth Map Wireframe Mesh Textured Model

Gauss-Seidel Relaxation. Update depth estimate at each pixel based on measured surface normals and previously estimated values for pixel's neighbors in the depth map. Iterate.

iPhone as 3D Scanner

The same technology runs on an iPhone in the form of an app called Trimensional, which can also output any 3D scan to a 3D printer.

Real-time Reconstruction

Continually light the scene, capture an image, recompute surface normals, update depth map, and display resulting textured model.

Currently runs at 10 frames per second on 320x240 images. Timings for each step of the algorithm during a single cycle are: surface normal computation (9 ms), depth map computation (32 ms), combined lighting, 3D display, and video capture (26 ms).

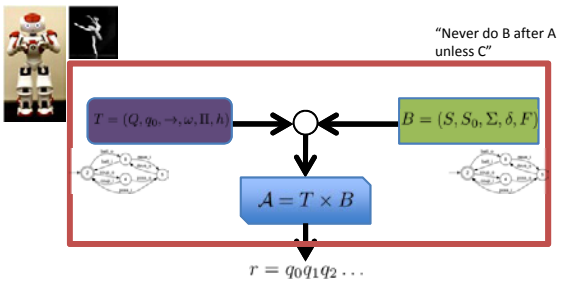
Capturing Style for Robotic Task Specification and Analysis of Aesthetic Movements

Amy LaViers and Magnus Egerstedt

Overview: Through analysis of formalized movement genres, namely, classical ballet, we have constructed a method for enumerating stylistic principles. With this quantified instantiation of style, we are able to assemble these principles as movement rules and automatically generate robotic motions which respect the desired stylistic task. Furthermore, such tools can enable quantitative study of style as it is employed in human behavior – through the creation of distinct movement genres such as ballet, hip-hop, and flamenco dance to name only a few.

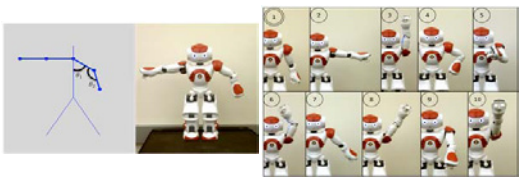
Here, we provide an overview of the work we have done in several publications. First, we look at capturing an essential essence of classical ballet in order to establish a reasonable method. Then, we apply this method to a real robotic platform. The sections below highlight key aspects of these works.

1) We can use an automata theoretic setting to combine the system dynamics with stylistic specifications. A transition system augmented with a set of atomic propositions enumerates allowable system transitions and which propositions are true at which state. The resulting product automaton accepts sequences of states which satisfy both the system dynamics and the specifications.



3) To apply this framework on an Aldebaran NAO robotic platform, we establish the following single arm states and dynamics. Using two different dynamics, we begin to establish two distinct stylistic behaviors: disco and cheerleading.

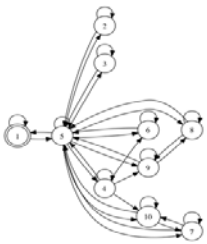
States for NAO arm dynamics:



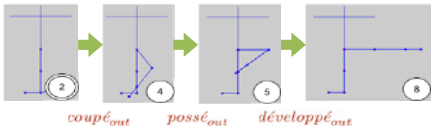
Allowable single arm transitions:

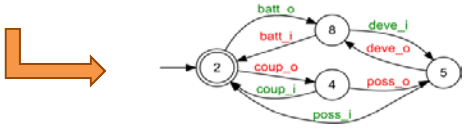
Cheerleading style

Disco style



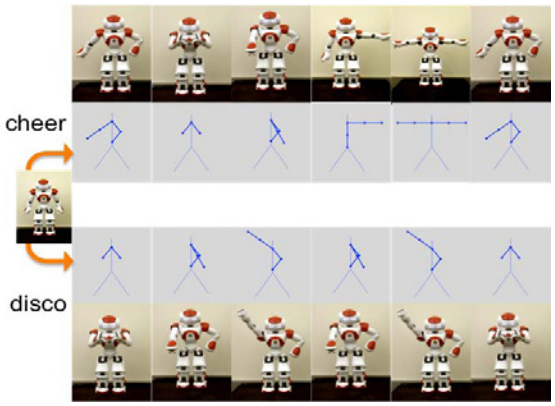
2) We deconstruct movements into two components: poses important to the experience of the movement and transitions between such goal poses. Hence, the model becomes a transition system where states correspond to body poses (usually angles of a single limb – leg shown below) and events correspond to allowable primary motions between these poses.





4) These two single arm systems are then combined to make a two arm system. This joint system allows both physically infeasible (i.e., when the arms collide) and aesthetically disallowed poses and motion sequences. We can restrict the behavior of these joint automata using formal composition methods. The specifications can be split into physically infeasible and aesthetically disallowed. In the case of the NAO, both styles will have the same physically infeasible restrictions since they take place on the same physical platform while each receives its own aesthetically driven restrictions specific to the desired style.

5) Snapshots of resulting sequences are shown below. Images of the robot accompany simulated abstractions of each pose.



Visit our website!: http://gritslab.ece.gatech.edu/ http://www.prism.gatech.edu/~alaviers3/

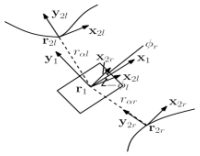
Simultaneous Cooperative Exploration and NeTworking (SCENT)

Jonghoek Kim, Sean Maxon, Fumin Zhang, and Magnus Egerstedt
Electrical and Computer Engineering, Georgia Institute of Technology, U.S.A.

Single Vehicle Exploration

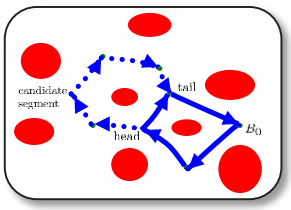
Tracking Control Law

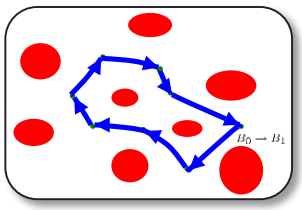
$$u = \frac{1}{2}\left(\frac{\kappa_r \cos(\phi_r)}{1-\kappa_r r_{\alpha r}} + \frac{\kappa_l \cos(\phi_l)}{1+\kappa_l r_{\alpha l}}\right) + 2\lambda(r_{\alpha l} - r_{\alpha r})(\cos(\phi_l) + \cos(\phi_r)) + \mu \sin\left(\frac{\phi_l + \phi_r}{2}\right).$$



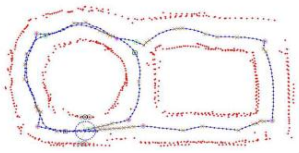
The Boundary Expansion Algorithms

Construct the Initial Enclosing Boundary. ⇒ Expand the Enclosing Boundary.

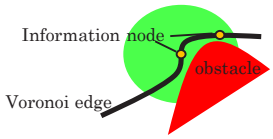


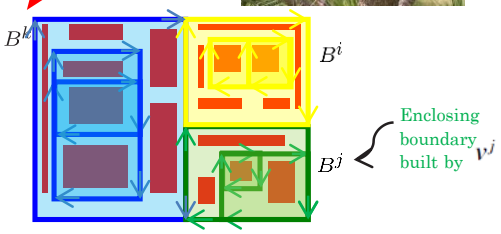


- The BE algorithms are provably complete.
- To implement the BE algorithms, only a circularly linked list is stored and updated based on two simple rules.

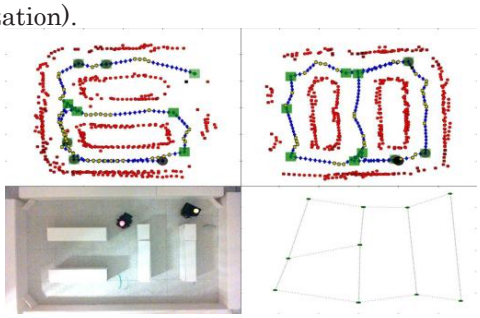


Multiple Vehicles Exploration



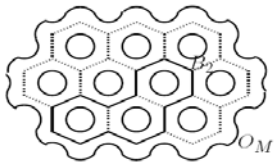


- Each vehicle deploys information nodes while expanding the explored area using the BE algorithms.
- Subgraphs laid out by different vehicles are merged into one maximally connected graph when they intersect.
- Coordinate of an information node is not necessary (topological localization).

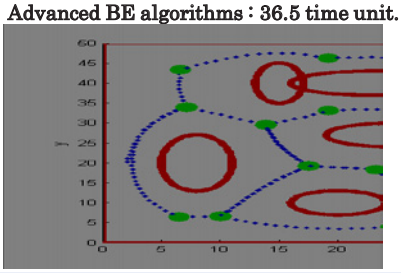


Exploration Time Analysis

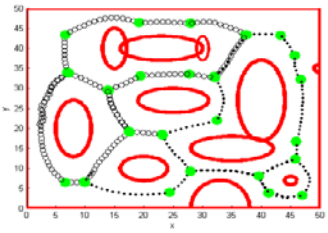
- There are M obstacles in the workspace, and all Voronoi cells are hexagonal with identical size.
- The exploration time using one vehicle is $O(M^2)$
- The exploration time using N vehicles is upper bounded by $a + \lceil \frac{M-1}{N} \rceil b + (\lceil \frac{M-1}{N} \rceil)^2 c$



Advanced BE algorithms : 36.5 time unit.



Advanced SCENT algorithms : 17.35 time unit.



References

J. Kim, F. Zhang, and M. Egerstedt. A Provably Complete Exploration Strategy by Constructing Voronoi Diagrams, Autonomous Robots, Vol. 29, No 3-4, pp. 367-380, 2010.

J. Kim, F. Zhang, and M. Egerstedt. Simultaneous Cooperative Exploration and Networking based on Voronoi diagrams, in Proc. of 2009 IFAC Workshop on Networked Robotics (NETROBO 2009), 1-6.

J. Kim, F. Zhang, and M. Egerstedt. An exploration strategy based on construction of Voronoi diagrams with provable completeness, in Proc. of 48th IEEE Conference on Decision and Control (CDC 2009), 7024-7029.

J. Kim, S. Maxon, F. Zhang, and M. Egerstedt. Simultaneous Cooperative Exploration and Networking based on Voronoi diagrams, submitted to IEEE Transactions on Robotics.

Georgia Robotics and Intelligent Systems Laboratory

http://gritslab.ece.gatech.edu

Pancakes: A Software Framework for Robot and Sensor Network Applications

Patrick Martin, Jean-Pierre de la Croix, and Magnus Egerstedt

Abstract

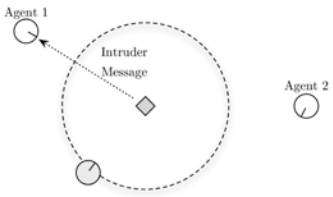
The development of control applications for multi-agent robot and sensor networks is complicated by the heterogeneous nature of the systems involved, as well as their physical capabilities (or limitations). We propose a software framework that unifies these networked systems, thus facilitating the development of multi-agent control across multiple platforms and application domains. This framework addresses the need for these systems to dynamically adjust their actuating, sensing, and networking capabilities based on physical constraints, such as power levels. Furthermore, it allows for sensing and control algorithms to migrate to different platforms, which gives multi-agent control application designers the ability to adjust sensing and control as the network evolves.

Application

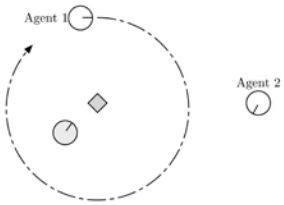
Building Security

Q: "Can we deploy a robot and sensor network application to detain an intruder and have it stay in operation?"

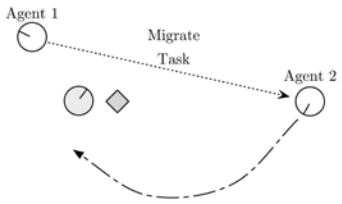
- Sensor network detects an intruder in the building and notifies the nearest robot response unit (Agent 1).



- Agent 1 responds and detains the intruder by encircling it.



- Agent 1 is low on power and delegates the task to a nearby patrolling unit (Agent 2). Agent 2 detains the intruder.



A: "Yes, we can with Pancakes."

Framework

Information Stream

"A configurable set of communication channels between software components."

- Serve as input and output channels.
- May be used to communicate information internally or over the network.

Tasks

"Input/Output software components that are functional units."

- Can be event-driven, time-driven, or a combination of both modes.
- Use communication channels to exchange information with other tasks.

Services

"Controllable sets of tasks."

- Adjust runtime behavior of tasks, such as start, stop, reschedule, or migrate tasks.

Features

Dynamic Adjustment

"Tasks may be adjusted during runtime to respond to changes in the state of the software, the network, or the physical platform."

- For example, a task may be forced to reduce its network transmissions to save power on the physical platform.

Migration

"Tasks may be migrated to another compatible Pancakes platform."

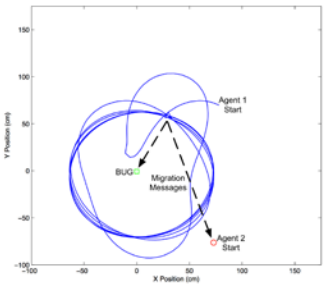
- Task can be migrated to another platform if it satisfies the task's dependencies.
- A cost metric is used to choose a platform for migration among candidates on a network.

Experiment

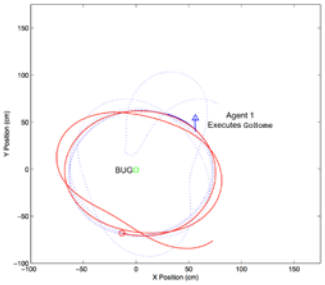
Building Security

"We deployed Khepera III mobile robots and BUG sensor nodes to detain intruders in our laboratory."

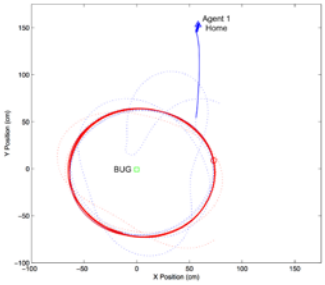
- A BUG sensor node detects an intruder in its region and notifies the nearest Khepera III mobile robot (Agent 1).
- Agent 1 responds and detains the intruder by encircling it.
- After some time, Agent 1 runs low on power and delegates the task to a nearby Khepera III mobile robot (Agent 2).



- Agent 2 accepts the migrated task and starts encircling the intruder. Agent 1 returns to its charging station.



- While Agent 1 has returned to the charing station, Agent 2 continues to detain the intruder.



The work by Patrick Martin, Jean-Pierre de la Croix, and Magnus Egerstedt was sponsored by the U.S. Office of Naval Research through MURI-HUNT.

Georgia Robotics and InTelligent Systems Laboratory

http://gritslab.ece.gatech.edu

Index-Free Multiagent Systems

An Eulerian Approach

Peter Kingston and Magnus Egerstedt

Abstract

Since the properties of systems comprising many homogeneous agents may be expected to be independent of how the agents happen to be indexed, it should be possible to formulate and solve multiagent control problems in an index-free way. In this paper we provide such an approach, based on an indicator distribution representation, which results in integro-differential dynamics that parallel and extend those obtained within the traditional indexed formulations. Conservation and stability properties are proven; a compatible geometric structure is constructed for the indexed representation; and a discrete analogue is presented which illustrates that for certain problems the Eulerian viewpoint results in very simple controllers.

Lagrangian Representation

"What state is each agent in?"

A Multiagent System

- N agents with states $x_1, \cdots, x_N \in \mathbb{R}^n$
- Collected in joint state vector $x = (x_1, \cdots, x_N) \in \mathbb{R}^{nN}$

Permutation Invariance

Equivalence relation:

$$x \sim y \iff \exists P \in \pi(N) \text{ s.t. } x = (P \otimes I)y$$

where $\pi(N)$ denotes the group of $N \times N$ permutation matrices. I.e., invariant under identity swaps:

$$(x_1, x_2, x_3) \sim (x_2, x_1, x_3)$$

Eulerian Representation

"How many agents are in each state?"

Indicator Distribution

$$m(x) = \Phi(x_1, \cdots, x_N)(x) = \sum_{i=1}^{N} \delta(x - x_i)$$

where δ is the Dirac delta distribution on $\mathbb{R}^n$, $\mathcal{T}(\mathbb{R}^n)$ denotes the space of tempered distributions on $\mathbb{R}^n$, and the map $\Phi : \mathbb{R}^{nN} \to \mathcal{T}(\mathbb{R}^n)$ creates m from $x_1, \cdots, x_N$.

Identical Controllers

In Lagrangian representation,

$$\dot{x}_i = v_i$$
$$v_i(t) = v(x_i(t), m(\cdot, t))$$

yields,

Advection equation

$$\dot{m} = -\operatorname{div}(mv) \qquad (1)$$
$$= -(\nabla m) \cdot v - m(\nabla \cdot v) \ .$$

Linear Consensus

Lagrangian setting

The joint state vector evolves according to

$$\dot{x}(t) = -L_w(\mathcal{G}(t))x(t)$$

where $L_w(\mathcal{G}(t))$ is the *graph Laplacian* for the interaction graph $\mathcal{G}(t)$.

Eulerian setting

The indicator distribution has the dynamics (1) with

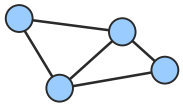
$$v = \int_{\zeta \in \mathbb{R}^n} m(\zeta) w(\zeta, x, m)(\zeta - x) d\zeta \ .$$

Properties of Index-Free Linear Consensus

- **Theorem 1:** Center of Mass Conservation
- **Theorem 2:** Stability

A Finite-State Analogue

A Graph of Rooms



- State: For each room i, a number $m_i \in \mathbb{R}$ of agents currently in that room.

Eulerian Dynamics

$$\left.\begin{array}{l} m[k+1] = m[k] + Du[k] \\ \text{s.t.} \quad m[k] \geq 0 \end{array}\right\} \forall k \in \mathbb{N}$$

where D is the incidence matrix associated with $\mathcal{G} = (V, E)$ and $u[k] \in \mathbb{R}^{|E|}$.

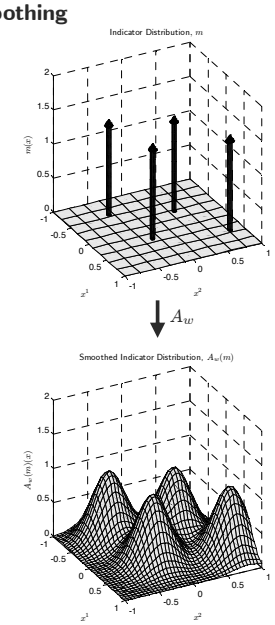
A Coverage Controller

$$u[k] = -\gamma D^T m[k] \ , \gamma > 0$$
$$\Downarrow$$
$$m[k+1] = [I - \gamma L(\mathcal{G})]m[k]$$

Inner Product Space via Smoothing

Imbue state space with geometric structure.

Smoothing



Kernelization

Compute directly with Lagrangian representation.

The *kernel* κ attached to the embedding $A_w \circ \Phi$ is the map

$$\kappa((x_1^1, \cdots, x_N^1), (x_1^2, \cdots, x_N^2)) \triangleq$$
$$\left\langle A_w(\Phi(x_1^1, \cdots, x_N^1)), A_w(\Phi(x_1^2, \cdots, x_N^2)) \right\rangle_{L^2}$$

which, for Gaussian smoothing,

$$= \sqrt{\frac{\pi^n}{\det Q}} \sum_{i,j} \exp\left(-\frac{1}{2}(x_j^2 - x_i^1)^T Q (x_j^2 - x_i^1)\right) .$$

Contributions

- Permutation-invariant representation.
- Integro-differential model.
- Properties of index-free consensus.
- Inner product structure and kernel.
- Explored Eulerian-Lagrangian / Coverage-consensus duality.

The work by Peter Kingston and Magnus Egerstedt was supported in part by the U.S. Office of Naval Research through MURI-HUNT.

Georgia Robotics and InTelligent Systems Laboratory
Georgia Institute of Technology
Robotics@GT & Intelligent Machines
GRITS
http://gritslab.ece.gatech.edu

Human-in-the-Loop: Terminal Constraint Receding Horizon Control with Human Inputs

Rahul Chipalkatty and Magnus Egerstedt

Abstract

This paper presents a control theoretic formulation and optimal control solution for integrating human control inputs subject to linear state constraints. The formulation utilizes a receding horizon optimal controller to update the control effort given the most recent state and human control input information. The novel solution to the corresponding finite horizon optimal control problem with terminal constraint is derived using Hilbert space methods. The control laws are applied to two planar human-driven mass-cart pendula, where the task is to synchronize the pendula's oscillations.

Introduction

- Humans are adept at high level tasks
- Automatic controllers capable for lower level tasks
- Shared Control/ Mixed Initiative / Human-in-the-Loop

Goal: Find a controller that preserves human intent while satisfying a linear constraint for the dynamic system

Problem Statement

- Given a linear discrete dynamic system

$$x_{k+1} = Ax_k + Bv_k$$

where v is the human input and x is the state at instant k

- Stay "close" to human input to preserve intent
- Satisfy a linear constraint at end of horizon
- Repeat for every instance we have new human input and state information
- Hence, a receding horizon optimal control formulation

$$\min_{\{u_k\}} V_N(\{v_k\},\{u_k\})$$

where

$$\begin{aligned} V_N(\{v_k\},\{u_k\}) &= \sum_{i=k}^{k+N-1} L(v_i,u_i), \\ L(v_i,u_i) &= (u_i-v_i)^T(u_i-v_i), \end{aligned}$$

subject to

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k, \\ x_{k+N} &\in \mathbb{X}_f = \{x \mid Mx = b\}, \end{aligned}$$

with $x_k \in \mathbb{R}^n$, $b \in \mathbb{R}^l$, and $u_k \in \mathbb{R}^m$.

- Pick N to weight human control/automatic control
- Predict where human is driving system
  - Currently estimating as a linear extrapolation
- Limits on human input, no limits on control input

Hilbert Space Solution

Find the control u that minimizes the cost V_N at time k

- Human input signal v in Hilbert space
- Constraint space V_α where inputs will drive system to satisfy linear constraints is an affine variety
- V_0 is the subspace
- Find the space orthogonal to the subspace and translate to pass through v
- Intersection is the control that minimizes the cost while driving to a terminal state that satisfies the constraint
- Control law at time instant k:

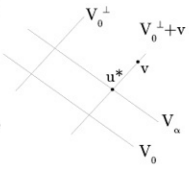


Figure 1: Hilbert's Projection Theorem

$$u_k^* = L_k^* M^T (MLL^* M^T)^{-1}(b - MA^N x_k - MLv) + v_k.$$

with

$$Lu = \textstyle\sum_{i=k}^{k+N-1} A^{k+N-1-i}Bu_i \qquad L_i^* = B^T(A^{k+N-1-i})^T$$

Receding Horizon Control Convergence

- Convergence results are well established in the literature.
- We adapt these proofs to show that indeed this control law will drive the state to the constraint set
- Assumptions:
  - Once the system is in the constraint set, the human is capable of keeping it there.
  - As a result, the cost is zero once in the constraint set.

Simulation

- Two Planar Mass-Cart Pendula
- States:
  - cart position and velocity
  - pendulum angle and angular velocity
- Human Input: Force commands to the cart
- Constraint:
  - Synchronize Swinging and Maintain cart distance

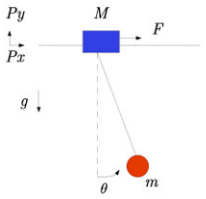


Figure 2: Mass-cart pendulum.

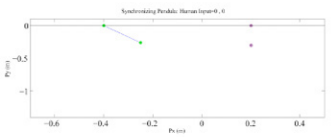


Figure 2: Graphic Interface of system: keyboard left right command 0.1 N increments of force to limits of -10N and 10N

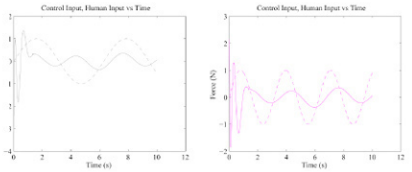


Figure 3: Human Input simulated as two sin waves with different frequencies. Hyperplane error goes to zero showing the linear constraint is satisfied.

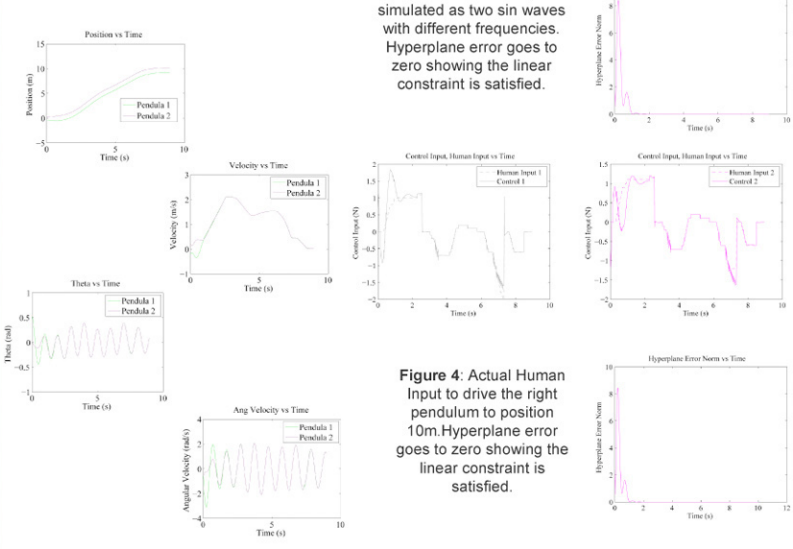


Figure 4: Actual Human Input to drive the right pendulum to position 10m.Hyperplane error goes to zero showing the linear constraint is satisfied.

Conclusions

- Formulation of a receding horizon optimal control problem to integrate human commands to drive a dynamic system to a linear state constraint
- Found a control law that solves this optimal control problem
- Proved this result converges to a state that satisfies the constraint
- Applied to a mass-cart synchronization problem to show the viability of the control law

Control of Multi-Agent Networks: From Network Design to Decentralized Coordination

Philip Twu *(ptwu@gatech.edu)* and **Magnus Egerstedt**
http://www.prism.gatech.edu/~ptwu3

Problem Statement

The research presented here addresses two key questions in the control of multi-agent networks.

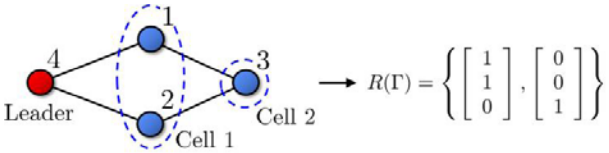
1. **How to design a multi-agent system that is appropriate for completing a given task?**
2. **Given a multi-agent system, how to design decentralized coordination strategies for it?**

Network Design: Homogeneous Single-Leader Networks

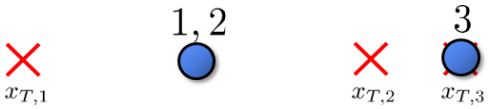
This project explores the **effects of agent diversity and network topology on a multi-agent system's expressiveness.**

In Single-Leader Networks:

- Expressiveness of network can be given by its controllability properties.
- Cells of maximal leader-invariant externally equitable partitions give the reachable subspace.



- **The network topology limits how close a target point can be reached.**



If Agents in the Network are Homogeneous:

- It does not matter which agent is assigned to which target point.
- Target labels can be permuted while still specifying the same desired configuration.
- **Some target point permutations are closer to the system's reachable subspace than others.**

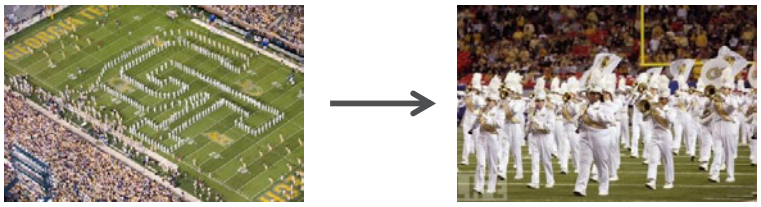
- **But, finding the optimal permutation involves solving a NP-hard problem.**

Publications:

- P. Twu, M. Egerstedt, and S. Martini, "Controllability of homogeneous single-leader networks," in *Proceedings of the 49th IEEE Conference on Decision and Control 2010*, vol. 2, pp. 1538–1543, 2010.

Optimal Decentralized Controller Generation

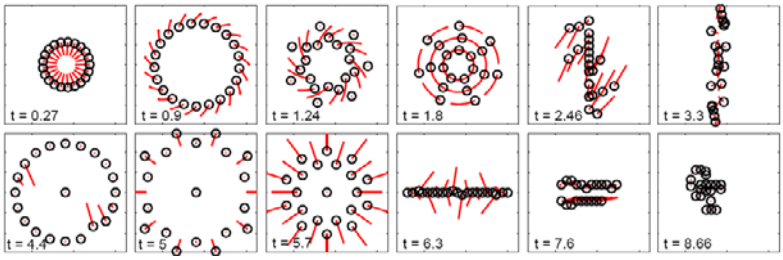
This project explores **how decentralized control laws for agents can be generated for tracking multi-agent motions.**



Overview of Approach:

1. **Specify parameterized constraints which describe what constitutes a decentralized controller for the system.**
2. Optimize the parameters to minimize the tracking error.

Results:



Publications:

- P. Twu and M. Egerstedt, "Optimal decentralization of multi-agent motions," in *American Control Conference (ACC), 2010*, pp. 2326–2331, IEEE, 2010.

Distributed Air Traffic Scheduling Algorithm

This project explores **using distributed negotiation algorithms to resolve terminal approach air traffic merging conflicts.**

Highlights of Algorithm:

- Opposing aircraft negotiate for:
  - Order of merging
  - Arrival time at merge point,
- while minimizing:
  - Fuel Consumption
  - Path Deviations
  - Changes in ETA.
- **Sufficient feasibility conditions proven for safe merging.**

Publications:

- P. Twu, R. Chipalkatty, A. Rahmani, M. Egerstedt, and R. Young, "Air traffic maximization for the terminal phase of flight under FAA's NextGen framework," in *Digital Avionics Systems Conference (DASC), 2010 IEEE/AIAA 29th*, pp. 2.C.1–1 –2.C.1–14, Oct. 2010.

Rockwell Collins
ONR Science & Technology

A Conceptual Space Architecture for Widely Heterogeneous Robotic Systems

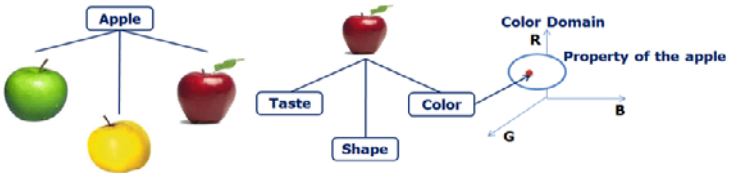
HyunRyong Jung, Arjun Menon, and Ronald C. Arkin
Mobile Robot Laboratory, School of Interactive Computing

Motivation

- Sharing knowledge across widely disparate robotic platforms is challenging.
- Robots are equipped with radically different sensors.
- Classical knowledge representations (e.g., symbolic representation and connectionist) have several deficits.

Conceptual Spaces

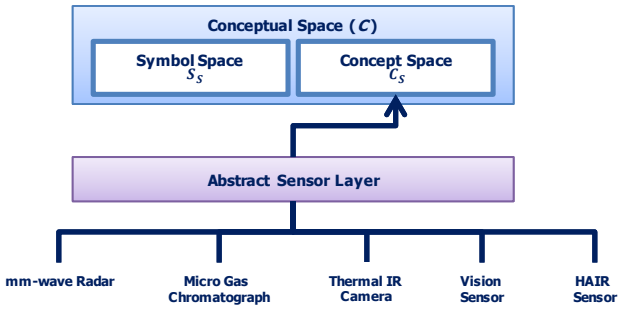
- Gärdenfors suggested the conceptual space, a metric world.
- Objects and abstract concepts are represented by quality dimensions.

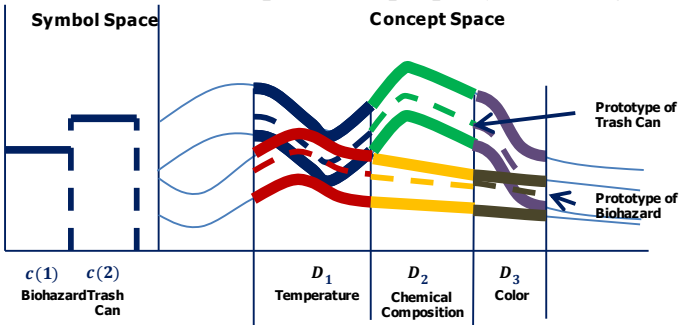


- The Similarity which is quite important in learning and induction, can be measured easily.

$$c(k) = \sum_{i=1}^{n} \alpha(k,i) \cdot s\big(p(k,i), f(k,i)\big)$$

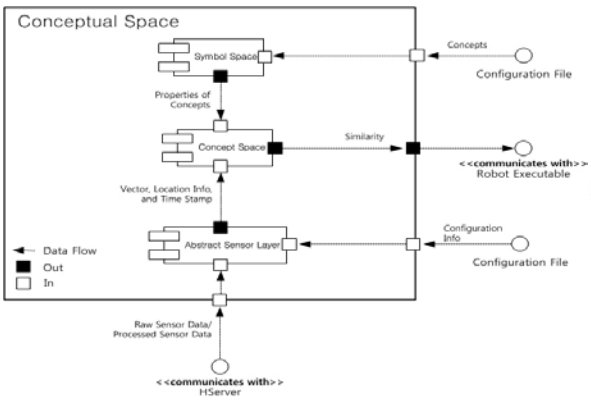
- Gaussian mixture model (GMM) for representing the region of a property
- Abstract sensor layer is to convert sensor data to vectors which can represent a property of an object.

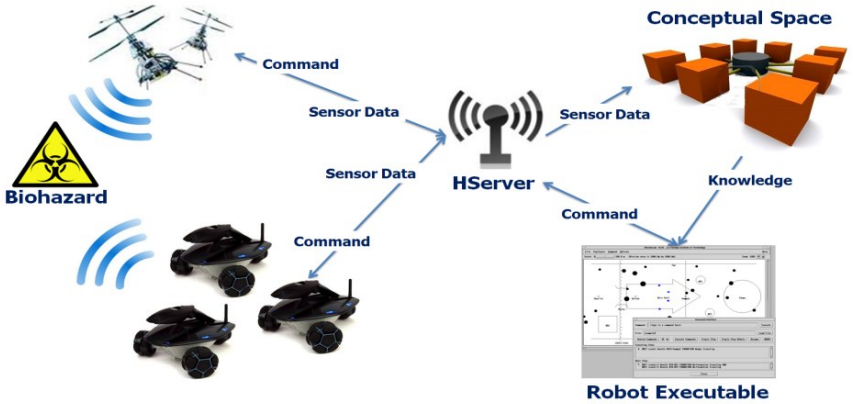




Architecture & Scenario

- Detect a biological weapon with two different robots: crawler and flyer.
- Properties of a biohazard - Barrel shaped, Chemical signature, Stored at low temperatures





HUNT: Heterogeneous Unmanned Networked Teams

- Future Naval Combat Operations and Systems will entail small expeditionary forces which must monitor and protect large and complex areas continuously
- The purpose of HUNT is to push the state-of-the-art in complex, time-critical mission planning and execution for large numbers of heterogeneous vehicles collaborating with warfighters
- Sophisticated cooperation among intelligent biological organisms will offer critical insight and solution templates for many hard engineering problems

Prototypical Progression of a Hunt

- Prototypical progression through foraging states shown to the right
- Hunt begins with wolves searching for prey
- Once prey is discovered the wolves transition to approach
- Transition to attack group when prey begins running
- Transition to attack individual when a weak individual is discovered
- Transition to capture when the prey is close enough to make contact
- Capture ends in a kill for successful hunts

Implementing Wolf Behavior

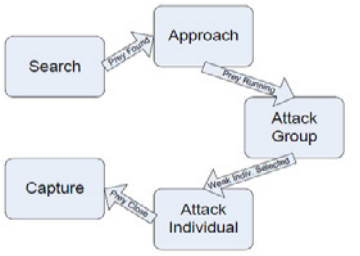
- Software implementation was accomplished through a set of releasers (stimuli) and a weighted roulette wheel of probabilities
- The presence or absence of stimuli make a transition possible (we say they release that transition)
- The set of releasers and the transitions they facilitate are shown in the table to the right

Finite State Automata

- The foraging states and transitions are represented by states and triggers in the Finite State Automata (FSA) shown below
- The FSA is a fully connected graph, however the only transitions possible are those with the necessary releasers present

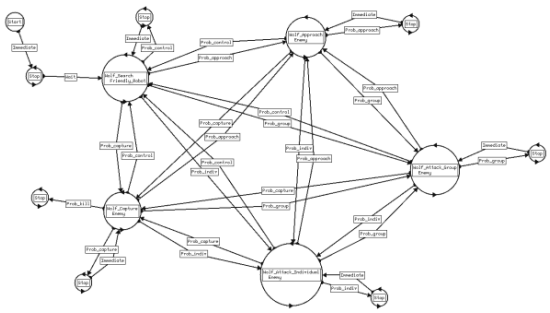
Results

- Experiments were run for 4 scenarios:
  - 1 wolf, 1 elk (stop)
  - 1 wolf, 1 elk (run away)
  - 1 wolf, 3 elk (run away)
  - 2 wolf, 3 elk (run away)
- Each scenario was run 20 times and the progression of the wolf through the foraging states was recorded for each run
- The tabulated values show a high fidelity to the original observed probabilities
- Differences between probabilities in each scenario show that the wolves behavior reacted to prey behavior



List of Releasers and Transitions

Releaser	Transitions possible
Prey Found	S → A, G, I
Prey Lost	A, G, I, C → S
Multiple Prey	S, A, I, C → G
Prey Running	A → G, I
Prey Stopped	G, I → A ; A, I → C
Weak Individual Identified	G → I
Prey Close	G, I → C



Robotics@GT & Intelligent Machines

Robots for Humanity

An accessible interface for mobile manipulation by the motor impaired

Healthcare Robotics
www.healthcare-robotics.com

Phillip Grice, Tiffany Chen, Hai Nguyen, Charles Kemp
Healthcare Robotics Laboratory, Georgia Institute of Technology

Georgia Institute of Technology

Purpose: To produce an accessible interface which allows persons with severe motor disability to control a fully articulated humanoid robot in performing self-assistive healthcare tasks.

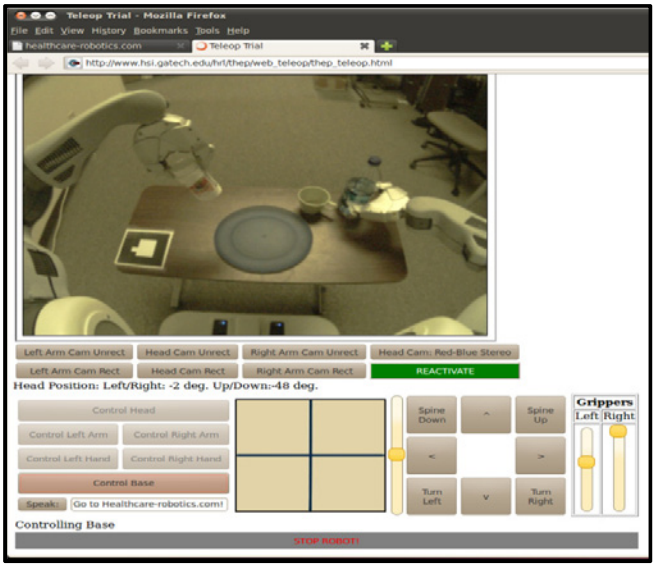
Overview:

This project involves a rapidly iterating, user-integrated design process to produce a user interface and underlying autonomous capabilities, allowing a single quadriplegic user/collaborator to physically interact with his environment through a PR2 robot.

In addition, the user provides a case-study in the needs of the severely disabled with respect to potential robotic solutions. His insight into patient needs and capabilities alike allows for the design of an interface which is both effective and easy to use, and remarkably powerful as a results

Interface:

- Web-based: **Only requires a browser**
- Direct control of all 25 physical DoF's
- Arm movements decomposed into:
  - 3D gripper position
  - Wrist joint configuration.
- Visual feedback from cameras in head and arms
- Text-to-speech
- Low-gain control of compliant, gravity compensated arms – **Safe for contact**
- Roll-over emergency stop to halt all motors.
- Designed from user-specified layout and developed with extensive, frequent user testing and feedback.



Initial Outcomes:

The user has used the interface and robot to:

- Perform remote manipulation tasks
  - Using the GT PR2 from home in California
- Interact and talk with others remotely
- Perform in-person object manipulation
- *Scratch his own face*
- *Brush his own hair and scratch his own head*

Significant need-finding has shown:

- Clear opportunities for Mobile Manipulators
  - Environmental control
  - Many ADL's
- A focus on communication
  - Currently slow, time-consuming
  - Potential for physical social presence

The user seated next to a robot under his control.

With the Web-based interface:
- No software is required
- Remote access is simplified
- Numerous commercial devices allow mouse control for a wide range of disabilities, which can now control the robot

Continuing Goals:

1. **Increase autonomy**
2. **Reduce required effort**

This will include:

- Visualization of 3D sensor data
- Choosing a point on a map for autonomous navigation.
- Point-and-click autonomous grasping and placement (Willow Garage)
- 'Perpendicular Approach' movement primitive
- Improved low-level control

Acknowledgements: The presenters would like to thank Mr. and Mrs. Henry Evans for their contributions and inspiration, and to recognize collaborative contributions and financial support from Willow Garage.

hsi Health Systems Institute
A Georgia Tech / Emory Initiative

Willow Garage

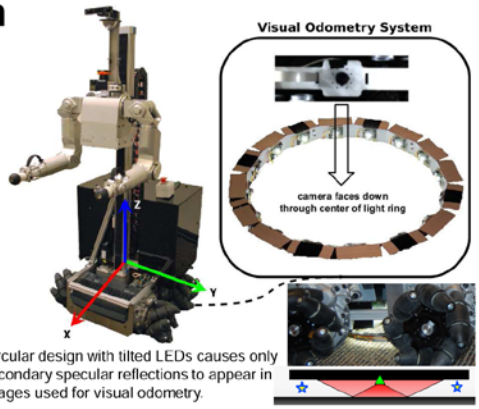
Robotics@GT & Intelligent Machines

Visual Odometry and Control for an Omnidirectional Mobile Robot with a Downward-Facing Camera

Marc Killpack, Travis Deyle,
Cressel Anderson, Charles C. Kemp

Motivation

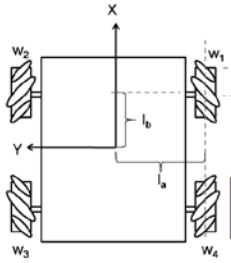
- Prior work in visual odometry has led to improvements over odometry from wheel encoders in cases with large amount of wheel slip
- Because our application domain is healthcare, we would like the system to be robust and function in common human environments (low lighting, crowded hallways, etc.)
- We believe that accurate control of our omnidirectional base will allow us to perform more complex mobile manipulation tasks



Circular design with tilted LEDs causes only secondary specular reflections to appear in images used for visual odometry.

Control of an Omnidirectional Robot

Off-axis forces and passive rollers make the mecanum type base a holonomic, omnidirectional platform



Inverse Actuated Solution[1]:

$$\begin{bmatrix}\omega_1\\ \omega_2\\ \omega_3\\ \omega_4\end{bmatrix} = \frac{1}{R}\begin{bmatrix}1 & 1 & l_{ab}\\ 1 & -1 & -l_{ab}\\ 1 & 1 & -l_{ab}\\ 1 & -1 & l_{ab}\end{bmatrix}\begin{bmatrix}V_x\\ V_y\\ V_\theta\end{bmatrix}$$

Sensed Forward Solution[1]:

$$\begin{bmatrix}V_x\\ V_y\\ V_\theta\end{bmatrix} = \frac{R}{4\,l_{ab}}\begin{bmatrix}l_{ab} & l_{ab} & l_{ab} & l_{ab}\\ l_{ab} & -l_{ab} & l_{ab} & -l_{ab}\\ 1 & -1 & -1 & 1\end{bmatrix}\begin{bmatrix}\omega_1\\ \omega_2\\ \omega_3\\ \omega_4\end{bmatrix}$$

Formulating the sensed forward solution into a discrete state space representation via a first order Euler approximation results in the following:

$$\begin{aligned} x(k+1) &= Gx(k) + Hu(k) \\ y(k) &= Cx(k) + Du(k) \end{aligned} \qquad u = [\omega_1\omega_2\omega_3\omega_4]^T$$

$$G = \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}, \quad H = \begin{bmatrix}tl_{ab} & tl_{ab} & tl_{ab} & tl_{ab}\\ tl_{ab} & -tl_{ab} & tl_{ab} & -tl_{ab}\\ t & -t & -t & t\end{bmatrix}$$

$$C = \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}, \quad D = [0]$$

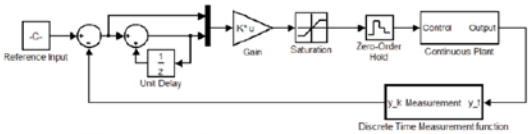
Doing a shift of variables and further augmenting the state allows us to formulate the state space with an integral term. This allows us to do integral control as can be seen below:

$$\hat{G} = \begin{bmatrix}G & 0\\ G & I\end{bmatrix}, \quad \hat{H} = \begin{bmatrix}H\\ H\end{bmatrix}$$

$$\hat{C} = \begin{bmatrix}C & 0\end{bmatrix}, \quad \hat{D} = 0$$

$$\left[x(\hat{k})\right] = \begin{bmatrix}x(k) - x(N)\\ \sum_{k=0}^{N}(x(k) - x(N))\end{bmatrix} = \begin{bmatrix}x_e(k)\\ \sum_{k=0}^{N} x_e(k)\end{bmatrix}$$

We implemented both a PID and LQR controller using the above kinematic state space models . Only the block diagram for the LQR controller is shown here:

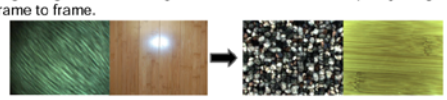


$$\hat{u}(k) = -K\hat{x}(k)$$

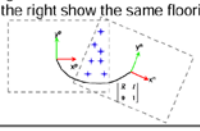
$$J = \frac{1}{2}\sum_{k=0}^{\infty}[\hat{x}(k)^T Q\hat{x}(k) + \rho\hat{u}(k)^T R\hat{u}(k)]$$

Visual Odometry Using a Downward-Facing Camera and LED Light Ring

Using the LED light ring shown in the figure above results in better quality images for feature tracking from frame to frame.



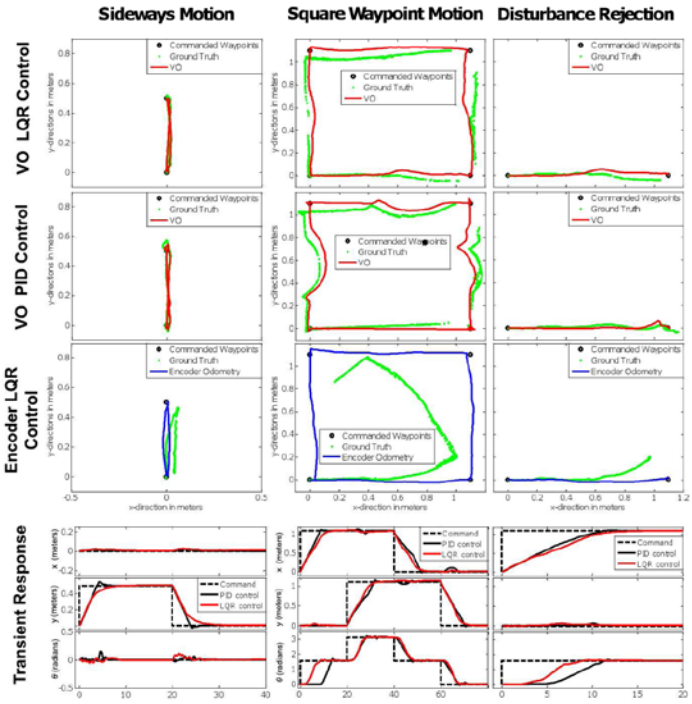
Images on the left show blur and reflection due to low light or a single light source, while images on the right show the same flooring while using the light ring.



Using Pyramidal Lucas-Kanade feature tracking and a homography, we estimate the change in x, y and theta in pixel units. We discard feature matches that are deemed as outliers from the initial solution and recompute a final differential odometry estimate before scaling it to meters.

Results of Representative Test Cases

We attached a large checker board to the robot and used an overhead camera to estimate the ground truth position of the robot.
We then performed three representative maneuvers useful in mobile manipulation as pictured to the right with the results shown below.



Contributions:

- Visual odometry estimates are sufficient to construct robust closed-loop PID and LQR controllers that perform significantly better (0.52% closed-loop error) than controllers using the wheel encoder odometry (14.4% closed-loop error)
- The odometry and control system is capable of operating in challenging conditions (e.g. large crowds or low ambient lighting)
- We have shown in other research that the system is robust enough for applications in difficult mobile manipulation tasks[2]:

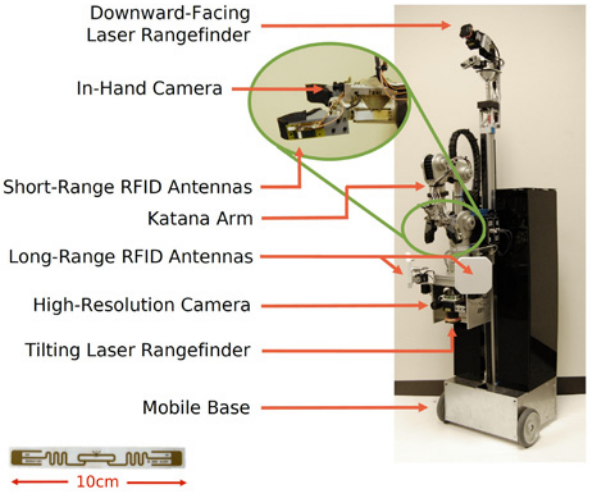
Healthcare Robotics | Robotics@GT & Intelligent Machines | Georgia Institute of Technology

[1] P. F. Muir and C. P. Neuman. "Kinematic modeling for feedback control of an omnidirectional wheeled mobile robot," In ICRA (1987)

[2] A. Jain and C. Kemp. "Pulling Open Doors and Drawers: Coordinating an Omni-directional Base and a Compliant Arm with Equilibrium Point Control." In ICRA (2010)

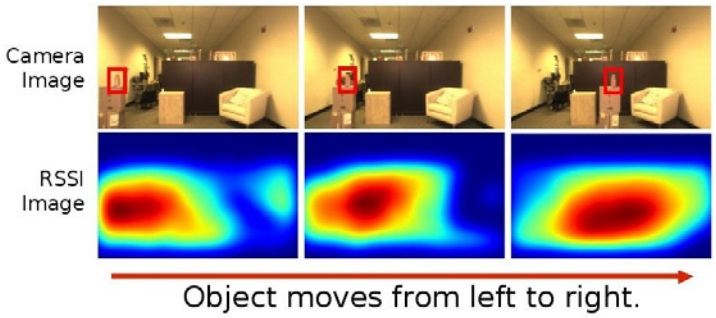
Ultra-High Frequency (UHF) Radio Frequency Identification (RFID) Sensing for Robotics Applications

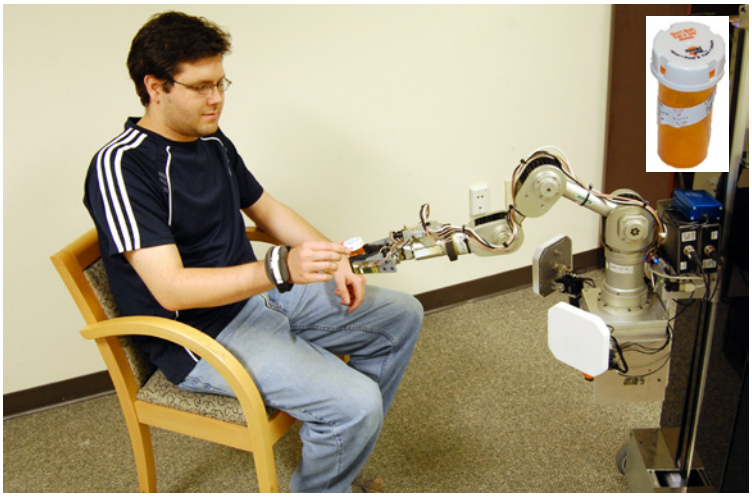
By: Travis Deyle & Prof. Charles C. Kemp
(Healthcare Robotics Lab)



The mobile manipulator "EL-E" has a variety of sensors and actuators. Of note, EL-E has two sets of UHF RFID antennas: A pair of long-range antennas can read tags up to 6 *meters* away; four short-range antennas embedded in EL-E's manipulator can verify the IDs of nearby or grasped objects. Both sets of antennas read the same UHF RFID tags, which are low-cost ($0.10), self-adhesive, and operate at ~900MHz.

We develop a number of UHF RFID sensing techniques and robot behaviors: tag localization, RSSI images (shown right), RFID servoing to approach tagged objects, multi-sensor fusion, semantic databases for robust applications, context-aware user interfaces, tag inventory behaviors, and grasped-object verification.





We develop a number of capabilities using this unique sensing modality. For example, robots that deliver the right meds to the right person at the right time could drastically improve medication adherence among older adults. EL-E locates and approaches a tagged person using long-range antennas, verifies the person and medication ID using short-range antennas, and then hands off the meds.

Robotics@GT & Intelligent Machines

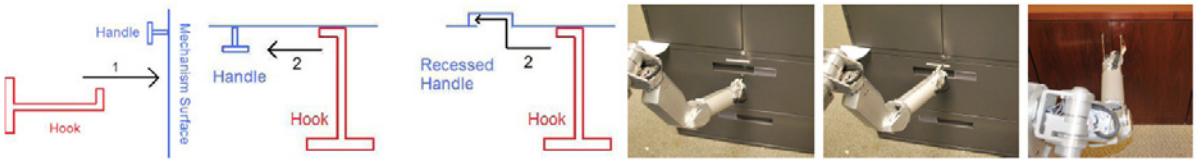
Leveraging Haptic Sensing and Low Mechanical Impedance for Mobile Manipulation in Unstructured Environments

Advait Jain Prof. Charles C. Kemp

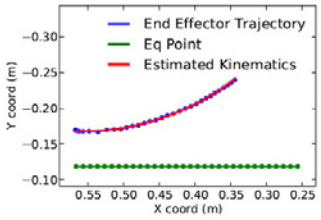
Healthcare Robotics Lab, Georgia Tech (www.healthcare-robotics.com)

Haptic Sensing to Reduce Uncertainty

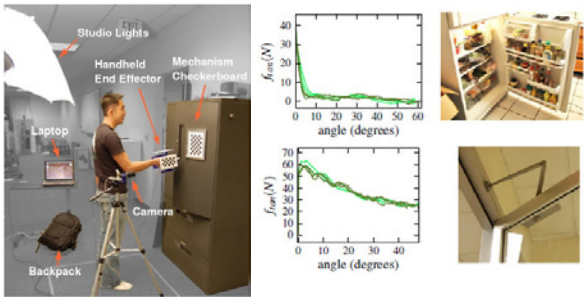
(a): TV Remote (b): Pill (c): Toothbrush (d): Cellphone (e): Keys (f): Pencil



Low Mechanical Impedance for Robust Performance

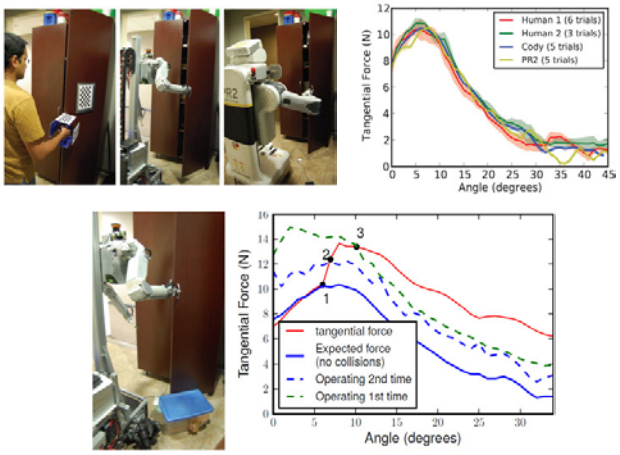


Capturing Everyday Mechanics



- Database of haptic common sense.
- Rational trade-offs in assistive robot design.

Sharing Haptic Experience



[1] A. Jain and C. C. Kemp. EL-E: An Assistive Mobile Manipulator that Autonomously Fetches Objects from Flat Surfaces. Autonomous Robots, 2010.
[2] A. Jain and C. Kemp, "Pulling Open Doors and Drawers: Coordinating an Omni-directional Base and a Compliant Arm with Equilibrium Point Control." ICRA, 2010
[3] A. Jain, H. Nguyen, M. Rath, J. Okerman, and C. C. Kemp. The Complex Structure of Simple Devices: A Survey of Trajectories and Forces that Open Doors and Drawers. IEEE RAS/EMBS BIOROB, 2010.
[4] A. Jain and C. C. Kemp. Sharing Haptic Experience for Improved Manipulation. IEEE Robotics and Automation Magazine (under review).

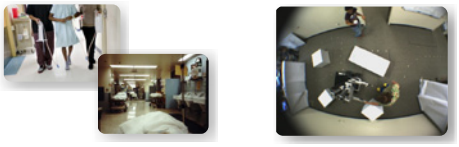
Lead Me by the Hand: Evaluation of a Direct Physical Interface for Nursing Assistant Robots

Tiffany L. Chen and Charles C. Kemp
Healthcare Robotics Laboratory | Georgia Institute of Technology

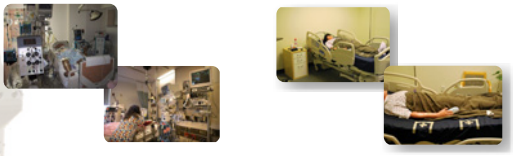
Motivation

- Older adult population is rising [1]
- Number of nurses and direct-care workers is decreasing [2]
- Robots may be able to assist nurses to perform care tasks such as bathing, feeding, and ambulating
- Using a direct physical interface (DPI) may provide an intuitive way for nurses to lead a robot to prepare for care tasks

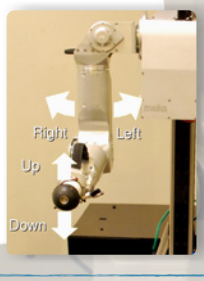
Healthcare Scenario 1: Hallways



Healthcare Scenario 2: Patient Room

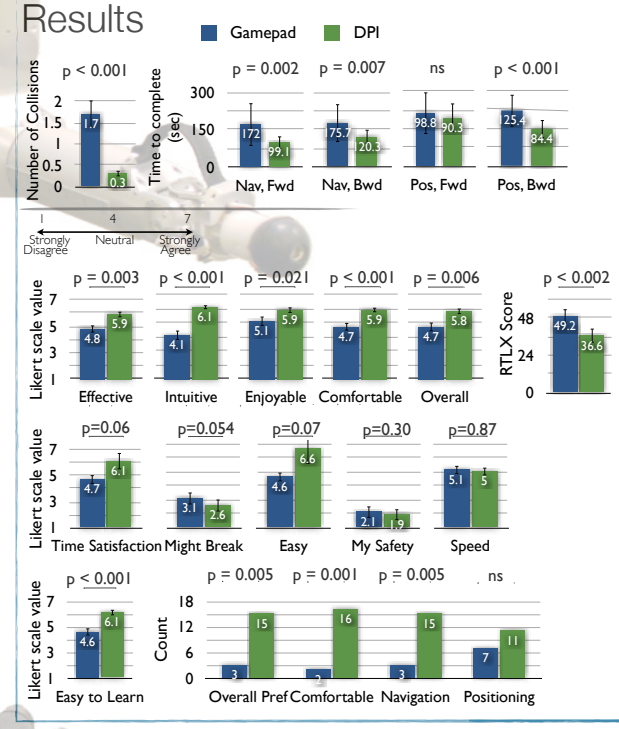


Direct Physical Interface



Gamepad Interface

Results



Experimental Design

- 18 nurse participants from Atlanta Area
- 2x4 Within-Subjects Design
- *Independent variables:*
  (1) Interface Type (DPI and Gamepad)
  (2) Task (Hallway Fwd, Hallway Bwd, Room Fwd, Room Bwd)
- *Dependent variables:*
  - Time to complete, # of collisions, NASA TLX, subjective ratings

Conclusions

- DPI found superior to comparable gamepad interface according to several objective and subjective measures
- DPIs may be well-suited for human-robot interaction (HRI) in nursing context and extend well to other fields

References

[1] Institute of Medicine. Retooling for an Aging America: Building the Health Care Workforce. The National Academies Press, 2008.
[2] H. J. Goodin. The nursing shortage in the united states of america: an integrative review of the literature. J Adv Nurs '03.

Acknowledgements

We would like to thank Advait Jain, Travis Deyle, Hye Yeon Nam, Kristina Falkenstrom, Aakanksha Gupta and all of the nurse participants. We gratefully acknowledge support from Hstar Technologies, the NSF Graduate Research Fellowship Program, Willow Garage, and NSF grant IIS-0705130.

Behavior Modeling & Turn-Taking Strategy For Robot Playmate

Hae Won Park and Ayanna M. Howard, Ph.D.
HumAnS Lab., School of Electrical and Computer Engineering

Background & Motivation

- ***My turn! Your turn!***
  - Interactive play helps promote cognitive, physical, and social skill developments [Piaget, 1952]
  - For children with developmental disabilities, *uniform* and *repetitive* exposure to interaction play is important [Dawson *et. al*, 2008]
- ***Previous Robot Playmates***
  - Help learn, diagnose, or treat cognitive disabiities
  - However, limited to non-autonomous imitation play
    - Kaspar / Robota / Tito [Dautenhahn 2000, Scassellati 2005, Billard 2006]

Algorithm

A. Sequencing Play Primitives

- Play Primitives
  - Basic motions that form manipulations when children interact with toys
  - Play behavior = Sequence of Play Primitives
  - Trained and recognized with 14 Hidden Markov Models
- Results
  - Average recognition rate: 86.88%

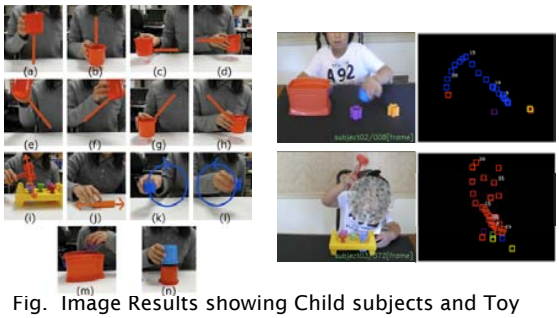


Fig. Image Results showing Child subjects and Toy Tracking for Play Primitive Sequencing

B. Planning Turn-Taking Strategy

- Case-Based Reasoning (CBR)
  - Solve new problems based on the solutions of similar past problems
- Child-Robot Turn-Taking CBR
  - Compare child's play to past play cases
  - Adapt to current play scene
- Results
  - Average recognition rate: 82.36%

	Play (a)	Play (b)
Problem (Play Scene)	Subject *stacks blue cup* on *orange cup*. Size ratio 0.85	Subject *inserts blue block* into *red bin*. Size ratio 0.11
Retrieve		
ttCBR retrieved Problem	stack blue cup on orange cup	insert blue block into red bin
ttCBR retrieved Solution	stack red cup on blue cup	insert purple block into red bin
Reuse		
ttCBR adapted Solution	stack {red cup} on blue cup	insert {orange,purple , green block} into red bin

Fig. Two play scenarios being evaluated by Turn-Taking CBR system & retrieving solution

Conclusion

- ***Modeling a play action by sequencing play primitives***
  - Decomposes a large action to a temporally-sequenced primitives with first order Markov process
  - Provides versatility to understand any kind of unpredicted behaviors
  - Achieves real-time recognition with low resolution camera
- ***Applying Case-Based Reasoning(CBR) for planning turn-taking task***
  - Solution is easily retrieved by comparing the child's play in the current scene to some past play cases stored in memory
  - Enables the whole system to bypass a repetitive long decision process

Visual Glacial Navigation

Stephen Williams and Ayanna M. Howard, Ph.D.
HumAnS Lab., School of Electrical and Computer Engineering

Background & Motivation

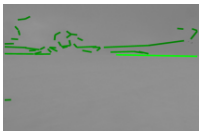
- ***Weather-related data from arctic regions can help scientists model climate change***
  - Automatic weather stations provide only spare coverage in Greenland and Antarctica
- ***Key technologies are needed to enable a robotic weather station network***
  - Detect and avoid arctic-specific hazards
  - Utilize of low-cost sensing technologies

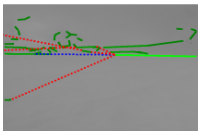
Horizon Detection

- ***"White-out" conditions make it difficult to separate ground from sky***
  - Low elevation clouds obscure background mountains
  - Overcast weather reduces available light and contrast

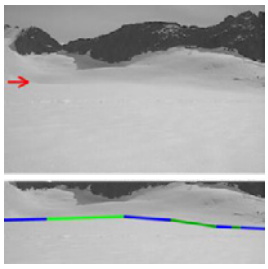
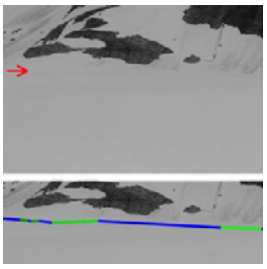

- ***Humans look for regions with strong information, then "fill in the gaps"***


Detect strong edges


Weight each edge using a heuristic set

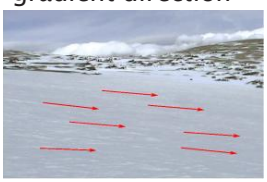

Perform a greedy search to connect segments

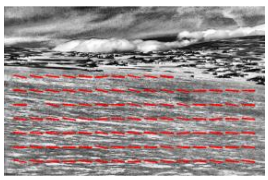
- ***Results***

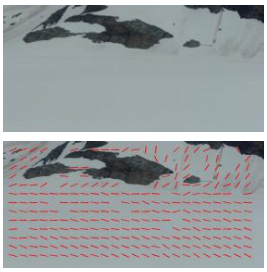
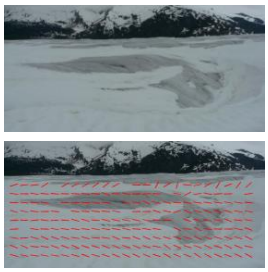
Slope Estimation

- ***Environmental hazards largely "slope-based"***
  - Few rocks/discrete obstacles
  - Gentle slope changes near mountains
  - Steep snow banks and lake basins
  - Cracks and crevasses
- ***Small-scale surface features and erosion patterns visually align***
  - Apply nonlinear contrast enhancement to emphasize small-scale texture
  - Perform least-squares estimate of dominant gradient direction


Original image showing perceived terrain slope


Enhanced image with slope estimates shown

- ***Results***

Behavior-based Control

- ***Distributed architecture integrates different, competing goals***
  - "Pursue Goal" and "Avoid Slope" behaviors implemented to minimize chassis roll

Trial Number	Chassis Roll Driving Directly To The Goal	Chassis Roll Using Ground Truth Slopes	Chassis Roll Using Visual Slope Estimates
1	11.07°	2.43°	2.55°
2	6.05°	3.58°	3.64°
3	3.99°	3.25°	3.30°
4	11.81°	5.32°	6.04°
5	8.93°	4.18°	4.51°

Conclusions

- ***Vision systems can extract useful terrain information, even in low-contrast glacial environments***
  - Methods developed for horizon detection, slope estimation, and hazard avoidance
  - Visual slope estimates produced similar navigational results to true slope data
- ***The low cost of camera systems is compatible with the requirements of multi-agent systems***

Science-centric Coverage Approaches for Intelligent Robot Navigation

Lonnie T. Parker and Ayanna M. Howard, Ph.D.
HumAnS Lab., School of Electrical and Computer Engineering

Background & Motivation

NASA (Earth Sciences Division) needs improved ground-based sensing of changes in environmental phenomena (i.e. soil moisture, chemical concentrations in air or water, and ice sheet deformation).

Previous research has shown that robotic technology can significantly aid the Earth scientist in this task, yet, the question remains: "*How should a robot navigate for the purpose of collecting data in an unknown environment?*"

We present a **robotic surveyor system** that incorporates on-line measurements to dictate the path of navigation. GOAL: Ensure that the area of interest is properly surveyed and minimizes error between estimated values and ground truth data.

Approach

Define the type of phenomena the robotic system will operate within (i.e. variable terrain, hills, valleys, etc.)

- 50 different DEMs generated to test navigation algorithms.

Design suitable navigation options for the chosen environment type .

- Data measured ONLINE should dictate the path that will provide the most dispersive sample set (i.e. best coverage) and yield the lowest resource usage (i.e. least distance traveled).
- Possible sampling options considered: Navigate according to a **Parallel Transect Framework** (i.e. Lawnmower), **Random waypoints**, or a custom **Piece-wise continuous** rule-set.
- Of these three, **Piece-wise continuous** is the only pattern that incrementally uses the measurements collected to dictate navigation *ONLINE*.

Execute each navigation option (in simulation), collecting samples along each path, and interpolate across the entire Area Under Test to create an estimated map for comparison with ground truth data.

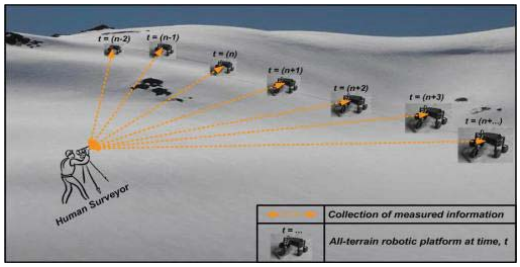


Fig 1. Future of robotic surveying – Teams of intelligent robot traversing the terrain and relaying relevant science data to a human surveyor

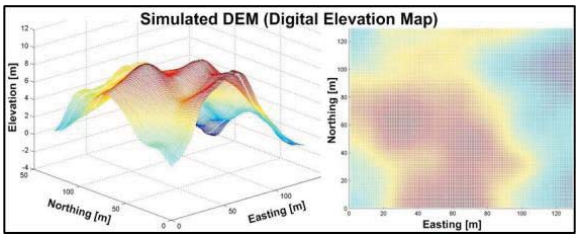


Fig 2. Sample digital elevation map

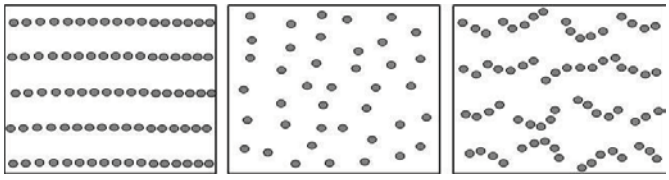


Fig 3. Sampling strategies that impact robot navigation

Results & Future Work

The **Piece-wise continuous** navigation rule-set proves to be the most beneficial, yielding the ***lowest RMSE*** across **50** simulated terrains and the ***lowest resource usage*** in terms of total Euclidean distance traveled.

The next step will be to develop navigation options for environmental phenomena other than elevation. Also, we will evaluate the performance of navigation patterns under noisy measurement conditions.

Quantifying Upper-Arm Rehabilitation Metrics for Children through Interaction with a Humanoid Robot

Douglas Brooks Ayanna Howard, Ph.D.
HumAnS Lab, School of Electrical and Computer Engineering

Manoi AT01

Problem

▪ Being able to objectively assess the performance of a patient through repeatable and quantifiable metrics has shown to be an effective means for rehabilitation therapy.

• However, to date, we are unaware of any research regarding child upper limb rehabilitation techniques using robotic systems; rather, the majority of these systems are only being applied to stroke patients.

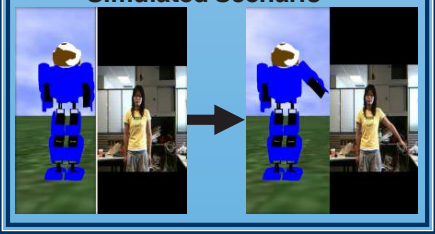
• While there are a number of robotic toys that have been shown to be engaging to children many of the studies focus solely on children with autism.

Goal

• The goal of this research is to fuse play and rehabilitation techniques using a robotic design to induce child-robot interaction that will be entertaining as well as effective for the child.

• We also want to use computer vision algorithms as the only means of evaluating proper rehabilitation techniques.

Simulated Scenario



Approach

Motion History Imaging

• Our algorithmic approach is to use temporal templates (MHI) to represent how motion in the image is moving.

• Using a replacement and decay operator, we obtain a scalar-valued image where more recently moving pixels are brighter.

$$H_\tau(x,y,t) = \begin{cases} \tau & \text{if } D(x,y,t) = 1 \\ max(0, H(x,y,t-1) - 1) & \text{otherwise} \end{cases}$$

Manhattan Distance

• Next, we divide the MHI's into grids, create feature vectors, and calculate the Manhattan Distance between two features vectors (the robot's and the patient's).

Dynamic Time Warping

• Using the distance calculations, we align image representations of robotic movements with those of the patient's movements.

• This allows us to determine whether or not the patient is doing the correct exercise.

$$dtw(m,n) = \begin{cases} W_I \cdot d(m,n) + dtw(m-1,n) \\ W_D \cdot d(m,n) + dtw(m,n-1) \\ W_S \cdot d(m,n) + dtw(m-1,n-1) \end{cases}$$

Contour Extraction & RANSAC

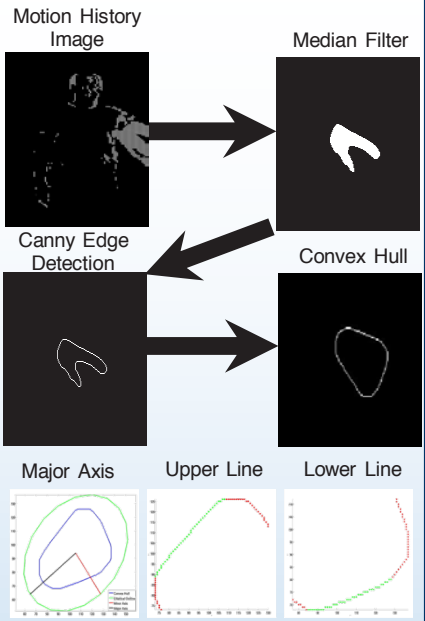
• We employ Median filtering, Canny Edge Detection, and a Convex Hull to extract a more ideal contour representing the patient's movement.

• Then we employ and elliptical fit and RANSAC in order to determine the upper and lower segments of movement.

Angular Velocity and Range of Motion

• Finally, we calculate Range of Motion and Peak Angular Velocity using the lines obtained from RANSAC and the associated timestamp.

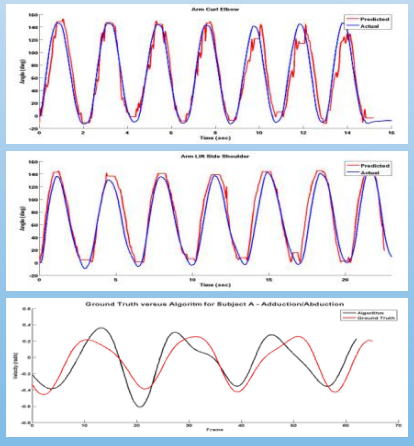
Resulting Image Sequence



Experimental Results

▪ We collected two types of exercise data from six different participants.

▪ ~ 92% Accuracy



Future Works

• Incorporate patient movement smoothness, stereo vision, equip robotic platform with software on-board, and test with child subjects.

Robotics@GT & Intelligent Machines

Incorporating a Model of Human Panic Behavior For Robotic-Based Emergency Evacuation

Paul Robinette and Ayanna M. Howard, Ph.D.
HumAnS Lab., School of Electrical and Computer Engineering

Background & Motivation

- ***Reduce Confusion***
  - Many avoidable injuries/deaths
  - Unfamiliar buildings
  - Unknown situation
- ***Why use robots?***
  - Reduce number of emergency personnel
  - Reduce response time
- ***Focus:***
  - **Guidance**
  - **Notification**
- ***Panic Models***
  - Close-knit (high affinity) groups stay together
  - Groups of strangers (low affinity) separate quickly
- ***Aircraft Evacuation***
  - Increased incentive *increased* time to evacuation
  - Evacuees very selfish

Methodology and Simulation

- 360° awareness
- 100 unit sight to lighted objects
- 50 unit sight to other humans
- 20 groups
  - Random size (m=5, σ=2)
  - Random affinity (high or low)
  - Random position (m=group center, σ=50)

HUMAN SIMULATIONS

Low Affinity Groups

Priority

Rules
Go to exit
Follow robot
Cluster with nearby humans
Panic

High Affinity Groups

Priority

Rules
Find group
Go to exit
Follow robot
Cluster with nearby humans
Panic

ROBOT SIMULATIONS

Priority

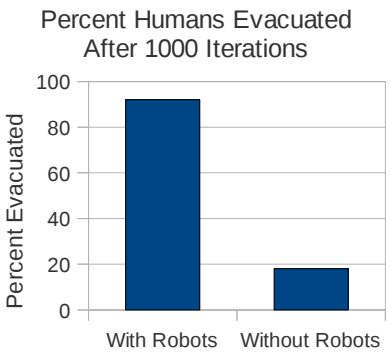
Rules
Go towards humans
Go towards exit
Explore

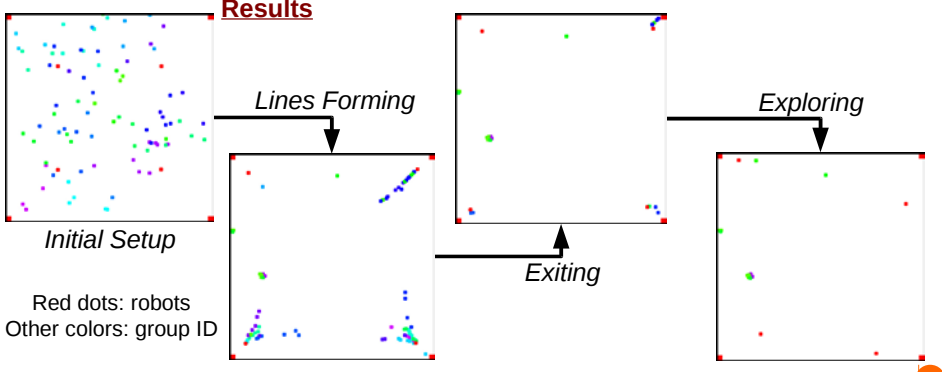
- 360° awareness
- 100 unit range to detect humans
- 50 unit range to lead humans

ENVIRONMENT

- Four exits in corners
- 500x500 units
- 1000 iterations

Results





Conclusion

- ***Robots enabled better evacuation***
  - More people evacuated within time limit
  - Orderly lines to exit
  - High and low affinity groups exited

Future Work

- **Use swarming/collaborative robot behavior**
- **Add third affinity level**
- **Add obstacles**
- **Create simulated human leaders**

Haptic Fusion of Multi-modal Interaction through a Mobile Manipulation Robotic System

Chun Hyuk Park and Ayanna M. Howard, Ph.D.
HumAnS Lab., School of Electrical and Computer Engineering

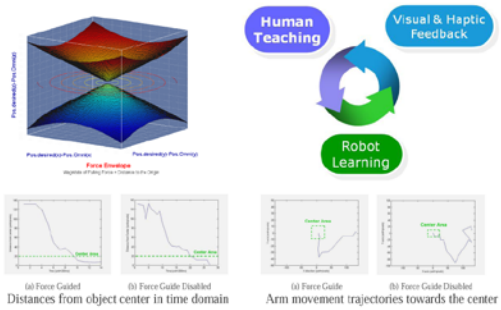
Background & Motivation

- ***Haptic modality for more effective Human-Robot-Interaction (HRI)***
  - ✓ High degree of controllability (Human → robot)
  - ✓ Increased feedback modality (Robot → Human)
  - ✓ Impact on learning process and the bandwidth of communication in HRI (Human ↔ Robot)
- ***Mobile manipulator as a multi-purpose assistive robotic platform***
  - ✓ Capable of aiding people with disabilities with multi-degree-of-freedom tasks
  - ✓ Integrated robotic sensory perception can be transferred to the human user to provide assistance and information for remote places

Approach and Result

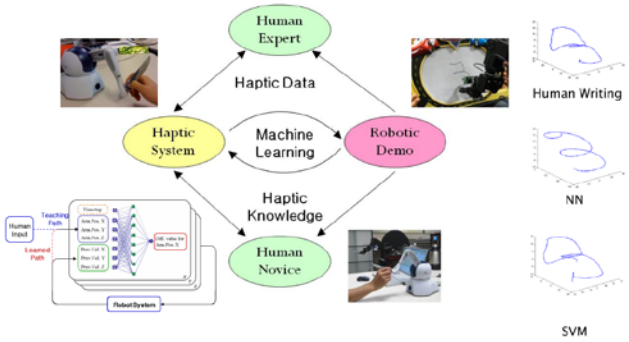
A. Haptic guidance

- Task-based force guidance for increased performance in both control and learning
- Human-in-the-loop interactive cycle



B. Haptic skill transfer

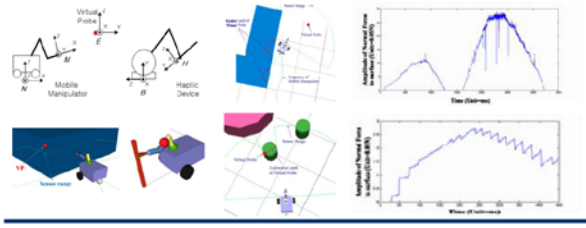
- Learning from the temporal sequences of sensing and actuation data
- Minimal training data



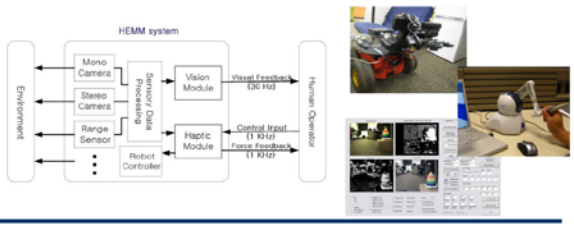
C. Haptic environmental feedback

- Integration of heterogeneous sensory perception
- Representation of the environmental knowledge with multi-modal feedback
- Point-cloud based haptic rendering

< Simulation >
- Sensor-based haptic feedback
- Increased safety and modality



< Hardware Implementation >
- Point-cloud-based haptic display
- Generated based on real-world environment



Conclusion & Future Work

- By utilizing heterogeneous sensors of a mobile manipulator, a human user can "feel" the environment before direct contact. Accordingly, easy operation and increased safety are achieved..
- Experiment results show that real-time haptic exploration of the environment adds a new dimension to haptic tele-operation and shows merits for mobile manipulation application in assistive technology.
- Next steps will be fusing multi-modal sensory perception (vision, sound, etc), designing the fusion model, and incorporating more accurate commercial sensors such as Xbox Kinect.
- More in-depth experiments with human subjects will be conducted.

Robotics@GT & Intelligent Machines

Combining Motion Planning and Optimization for Flexible Robot Manipulation

Jonathan Scholz & Mike Stilman, ICHR 2010

Motivation:

Robots face two challenges in natural environments:

- *Underspecified goals:* no human to specify exact goal configuration
- *Uncertain dynamics:* Effects of robot's actions on novel objects is uncertain

Approach:

- For underspecified goals:
  - *Pose task as a constrained optimization problem over a set of reward or cost terms.*
  - *Can be defined manually or modeled from human*
- For uncertain dynamics:
  - *Quickly approximate dynamics for a set of actions*
  - *Plan efficiently using sampling-based techniques*

Our algorithm:

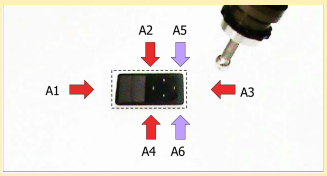
- Searches in *object* configuration space using Rapidly-exploring Random Trees (RRT)
- Adds leaves to search tree by forward-simulating the learned dynamics for each object-action pair
- Uses directGD heuristic to quickly search optimization landscape
- Returns a plan from the starting state to the most optimal reachable state, given cost function

```
TASK_SPACE_RRT(s_init, A)
 1  for i ← 1 to |O|
 2  do Model ← LEARN_MODEL(O_i, A, s_init)
 3  T.init(s_init)
 4  for i ← 1 to max_nodes
 5  do s_GD ← DIRECTGD(T)
 6     if RAND() > ε
 7        then s_samp ← s_GD
 8        else s_samp ← RANDOM_CONFIG()
 9     s_near ← NEAREST_NEIGHBOR(s_samp)
10     a* ← ARGMIN_a(ρ (MODEL(s_near, a), s_samp))
11     s_new ← MODEL(s_near, a*)
12     if not IN_COLLISION(s_new)
13        then ADD_VERTEX(s_new)
14             ADD_EDGE(s_near --a*--> s_new)
```

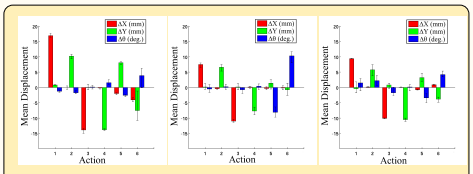
Pseudo code for TS-RRT algorithm

Model Learning

Goal: discover the dynamics of each object class over a set of action primitives



Six action primitives defined for interacting with objects in the robot's workspace



Model learning results from three object types, depicting mean and 95% CI for object displacement in each workspace dimension

```
LEARN_MODEL(o, A, s_init)
1  s = s_init
2  for i ← 1 to |A|
3  do while σ² > σ²_ref
4     do (Δs, s) = APPLY_ACTION(a_i, O)
5        P(Δs|a_i, o) = UPDATEDISTR(P(Δs|a_i, o), a_i, Δs)
6        σ² ← VARIANCE(P(Δs|a_i, o))
```

Model-learning procedure

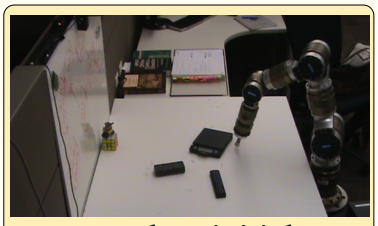
Manipulation under uncertainty

Initial state

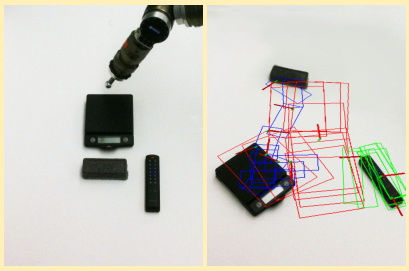
- Robot begins with a workspace containing three unfamiliar objects
- Robot provided a cost function expressing the following desiderata:
  - *Orthogonality*
  - *Cicumscribed area*
  - *Distance from edge of workspace*

Solution

- All objects pushed to orthonormal orientations in the center of the workspace
- All paths free of collisions and redundant actions
- Robot monitored error and replanned as necessary



Random initial configuration

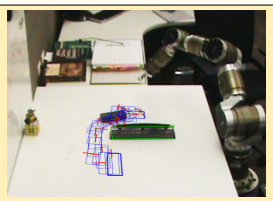


Optimized final configuration

Generalization to other manipulation tasks

Appropriate for tasks naturally expressed as optimization of a cost function:

- *Arranging clutter on a surface*
- *Multiple object placement*
- *Table setting*



In the presence of fixed obstacles

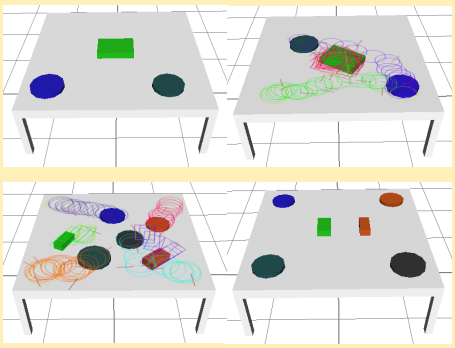


Table-setting with arbitrary numbers of objects

Advantages

Appropriate for tasks naturally expressed as optimization of a cost function:

- Unlike conventional single-shot methods, doesn't require user specified goals
- Always guaranteed to return reachable solution
- Favorable anytime characteristics
- Feasible for real-time planning in high DOF problems

Similar to Reinforcement Learning formalism, but trades path optimality for realtime feasibility

- RL can require many full-passes through configuration space to converge to optimal policy
- Handling continuous features requires discretization, tiling, or other appoaches
- RL better suited for problems with sparse reward landscape, but optimizations offer a gradient (like shaping reward) which allows fast heuristic search with RRT

The Motion Grammar for Physical Human-Robot Games

Neil Dantam, Pushkar Kolhe, Mike Stilman

Center for Robotics and Intelligent Machines, Department of Interactive Computing,
Georgia Institute of Technology, Atlanta, GA 30332

ntd@gatech.edu, pushkar@gatech.edu, mstilman@cc.gatech.edu

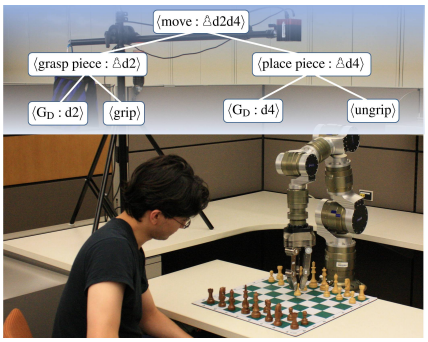


Figure 1: *Experimental setup for human-robot chess and partial parse-tree indicating the robot's plan to perform a chess move.*

Abstract

WE INTRODUCE THE MOTION GRAMMAR *[1], a powerful new representation for robot decision making, and validate its properties through the successful implementation of a* physical *human-robot game. The Motion Grammar is a formal tool for task decomposition and hybrid control in the presence of significant online uncertainty. This work describes the Motion Grammar, introduces some of the formal guarantees it can provide, and represents the entire game of human-robot chess through a single formal language. This language includes game-play, safe handling of human motion, uncertainty in piece positions, misplaced and collapsed pieces. We demonstrate the simple and effective language formulation through experiments on a 14-DOF manipulator interacting with 32 objects (chess pieces) and an unpredictable human adversary.*

1. The Motion Grammar

1.1 Overview

- A Linguistic Approach to Robot Control
- Discrete Events as Context-Free Grammar
- Continuous Semantics as Differential Equations
- Grammar used to produce the Motion Parser to control robot

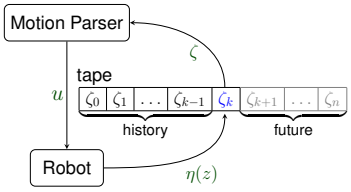


Figure 2: *Operation of the Motion Parser.*

1.2 Definition

Definition 1. *The Motion Grammar, $\mathcal{G}_M$, is a tuple* $\mathcal{G}_M = (\mathcal{Z}, Z, \mathcal{U}, \eta, V, P, K, S)$*:*

$\mathcal{Z}$ *space of robot sensor readings*

Z *set of tokens representing events*

$\mathcal{U}$ *space of robot inputs*

η *tokenizing function,* $\eta : \mathcal{Z} \mapsto Z$

V *set of nonterminals*

P *set of productions*

K *set of semantic rules, each associated with one and only one production.*

S *starting nonterminal,* $S \in V$

Definition 2. *The Motion Parser is a program that recognizes the language specified by the Motion Grammar and executes the corresponding semantic rules for each production.*

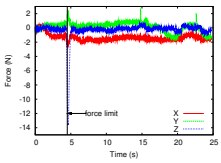
1.3 Benefits

- Immediate response to uncertainty and efficiency in the context of many degrees of freedom.
- Structure imposed by the grammar yields rigorous guarantees that controllers will respond to all contingencies.
- Hierarchical nature of grammatical productions and corresponding parse trees allows robot tasks to be recursively divided into simpler subtasks.

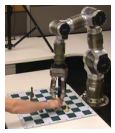
2. Guarded Trajectory Control

$$
\begin{array}{lcll}
\langle \mathrm{G} \rangle & \rightarrow & \langle \mathrm{G_D} \rangle \mid \langle \mathrm{G_L} \rangle & (1) \\
\langle \mathrm{G_D} \rangle & \rightarrow & \lfloor 1 \rfloor \mid \langle \kappa \rangle \langle \mathrm{G_D} \rangle & (2) \\
\langle \mathrm{G_L} \rangle & \rightarrow & \lfloor \mathrm{limit} \rfloor \mid \langle \kappa \rangle \langle \mathrm{G_L} \rangle & (3) \\
\langle \kappa \rangle & \rightarrow & \lfloor 0 \rfloor \ \{ \dot{\mathbf{q}} = \mathbf{J}^* \left(\dot{\mathbf{x}} - \mathbf{K}_p \left(\mathbf{x} - \mathbf{x}_r \right) - \mathbf{K}_f \left(\mathbf{F} - \mathbf{F}_r \right) \right) \} & (4)
\end{array}
$$

Figure 3: *Grammar fragment for guarded moves*



(a) Forces



(b) Contact

Figure 4: *Grammatical guarded moves safely protecting the human player.*

- Force limit at 4.7N, Fig. 4(a)
- Fast online control – responds immediately.
- Production (3) of Fig. 3 *guarantees* that when this situation occurs, the robot will stop.

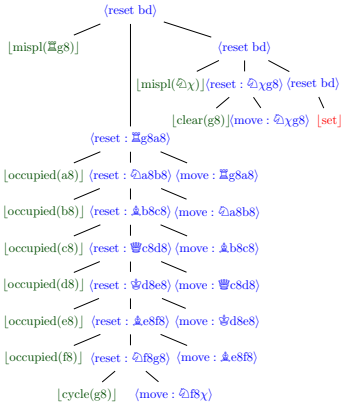
3. Board Resetting

$$
\begin{array}{lcl}
\langle \text{reset board} \rangle & \rightarrow & \lfloor \text{misplaced(x)} \rfloor \langle \text{reset : x, home(x)} \rangle \langle \text{reset board} \rangle \\
 & \mid & \lfloor \text{set} \rfloor \\
\langle \text{reset : } x_0, x_1 \rangle & \rightarrow & \lfloor \text{clear}(x_1) \rfloor \langle \text{move : } x_0, x_1 \rangle \\
 & \mid & \lfloor \text{occupied}(x_1) \rfloor \langle \text{reset : } x_1, \text{home}(x_1) \rangle \langle \text{move : } x_0, x_1 \rangle \\
 & \mid & \lfloor \text{cycle}(x_1) \rfloor \langle \text{move : } x_1, \text{rand}() \rangle
\end{array}
$$

Figure 5: *Grammar fragment to reset chessboard*

(a) Initial

(b) Final



(c) Parse Tree

Figure 6: *Example of board resetting*

- When occupied squares form a cycle, robot must break before resetting
- As the parser searches chain of pieces occupying each other's home squares, it builds up a stack of the moves to make
- Arbitrary depth search that is Context-Free – cannot be done with a Regular (finite state) Language

References

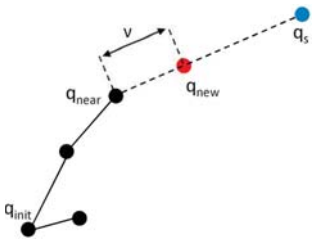
[1] N. Dantam, P. Kolhe, and M. Stilman, "The motion grammar for physical human-robot games," in *Proceedings of the IEEE International Conference on Robotics and Automation*. IEEE, 2011.

Sampling Heuristics for Optimal Motion Planning in High Dimensions

Baris Akgun and Mike Stilman
Center for Robotics and Intelligent Machines
The School of Interactive Computing
Georgia Institute of Technology

Motivation

- Robots are getting complicated and entering dynamic environments
- Optimality and efficiency of motion planning in high dimensions
- Iterative refinement of found plans in remaining planning time
- Advent of a probabilistically optimal planner (RRT*)



Rapidly Exploring Random Trees

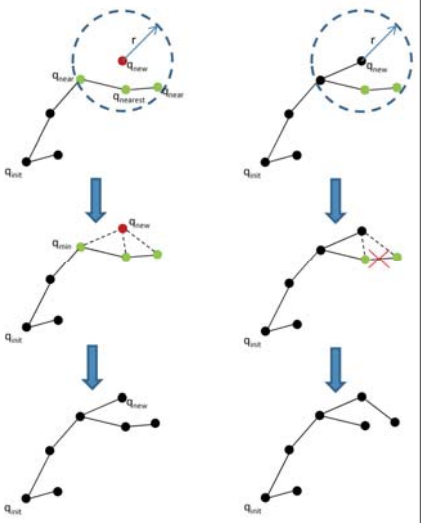
Typical Scenario

- High dimensional robot
- Limited planning time
- Dynamic environment

Approach

- Use RRT* to find an initial path
- Use developed heuristics to improve the iterative refinement process

RRT* Algorithm



Advantages of RRT*

1. Probabilistic Optimality
2. Iterative Refinement
3. Notion of cost in the tree

Challenges of RRT* in High Dimensions

1. Too much exploration
2. Path gets buried inside the tree
3. Not efficient

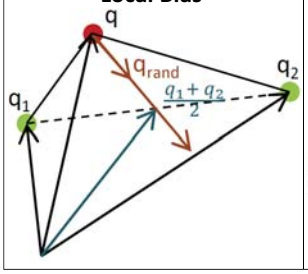
Heuristics

- Local-bias: Exploitation
- Node Rejection: Efficiency
- Bi-Directional: Efficiency

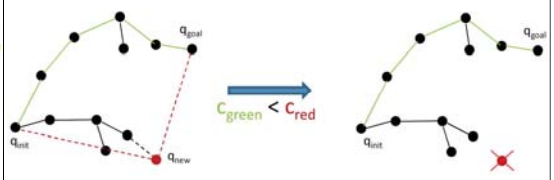
Resulting Planner

- Probabilistically Complete
- Probabilistically Optimal
- Efficient

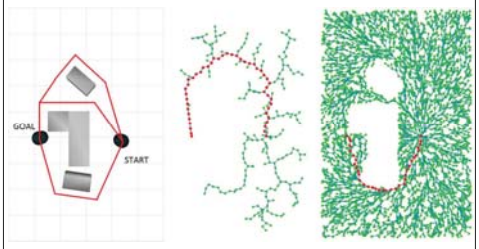
Local Bias



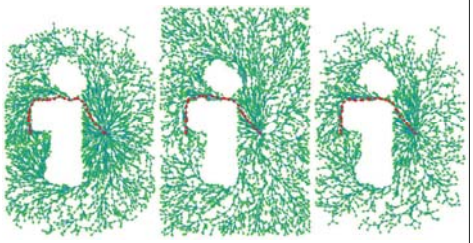
Node Rejection



2D Navigation Problem



Environment Initial Plain



Node Rejection Local Bias Combined

The environment and the resulting trees for the same initial path.

Results

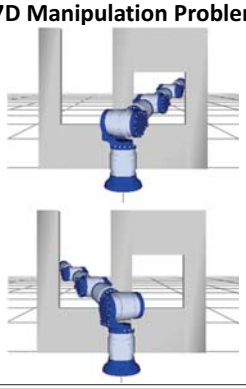
2D PERFORMANCE FOR A TIME BUDGET OF 1*sec*. RESULTS AVERAGED OVER 10 RUNS. PARENTHESES INDICATE STANDARD DEVIATION.

	c_1	c_{end}	Iterations
PL	8.07 (1.42)	5.84 (0.45)	5063 (166)
NR	8.07 (1.42)	5.16 (0.37)	9699 (1099)
LB	8.07 (1.42)	5.58 (0.68)	4732 (215)
CO	8.07 (1.42)	5.39 (0.71)	5593 (380)

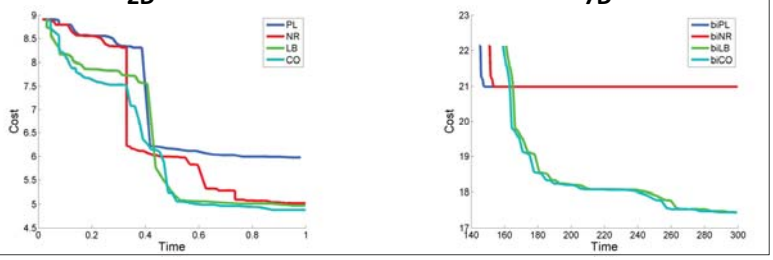
7D PERFORMANCE FOR A TIME BUDGET OF 300*sec*. RESULTS AVERAGED OVER 10 RUNS. PARENTHESES INDICATE STANDARD DEVIATION.

	c_1	c_{end}	Iteration
biPL	22.48 (1.86)	21.01 (1.39)	61766 (8869)
biNR	22.48 (1.86)	21.01 (1.39)	60043 (8942)
biLB	22.48 (1.86)	17.25 (1.66)	49991 (12359)
biCO	22.48 (1.86)	17.21 (1.66)	50209 (12384)

7D Manipulation Problem



Cost Evolutions



This work has been submitted to IROS 2011

Optimized Whole-body Motion Control for Wheeled Humanoid Robots

Kasemsit Teeyapan and Mike Stilman
kasemsit@gatech.edu, mstilman@cc.gatech.edu
Humanoid Robotics Lab, Georgia Tech

Robotics@GT & Intelligent Machines

Abstract

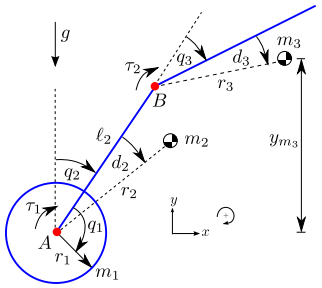
In this paper, the whole-body motion control of a dynamically-stable two-wheeled humanoid robot is investigated. With feedback linearization as a control law cooperating with a stochastic optimization technique called particle swarm optimization (PSO) to search for an optimal set of certain unknown parameters, an optimal robot motion can be realized according to a pre-fined performance index. The main contribution of this work is our first success on implementing the proposed control strategy on the actual physical system, Golem Krang, a novel two-wheeled humanoid robot developed at Georgia Tech. The experiments are demonstrated by two primitive motions, standing from the ground and deceleration.

Dynamics Model

The dynamics model of the robot is derived by Lagrangian mechanics as a two-wheeled planar double inverted pendulum resulted in the equation of the form

$$\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{V}(\mathbf{q}, \dot{\mathbf{q}}) = \tau, \tag{1}$$

where $\mathbf{q} = [q_1, q_2, q_3]^T$ and $\tau = [\tau_1, 0, \tau_2]^T$.



Control

Let $q_a = q_1 + q_2$. Differentiate two CoM equations twice, we obtain the following equations :

$$w_1\ddot{q}_1 + w_2\ddot{q}_2 + w_3\ddot{q}_3 + v_4 = \ddot{y}_{m_3} \tag{2}$$
$$w_4\ddot{q}_1 + w_5\ddot{q}_2 + w_6\ddot{q}_3 + v_5 = \ddot{q}_a \tag{3}$$

Augment them into Eq. 1, we get

$$\left[\begin{array}{ccc|cc} & \mathbf{M}(\mathbf{q}) & & -1 & 0 \\ & & & 0 & 0 \\ & & & 0 & -1 \\ \hline w_1 & w_2 & w_3 & 0 & 0 \\ w_4 & w_5 & w_6 & 0 & 0 \end{array}\right] \left[\begin{array}{c} \ddot{\mathbf{q}} \\ \hline \tau_1 \\ \tau_2 \end{array}\right] = \left[\begin{array}{c} \mathbf{V}(\mathbf{q}, \dot{\mathbf{q}}) \\ \hline \dot{y}_{m_3} - v_4 \\ \ddot{q}_a - v_5 \end{array}\right] \tag{4}$$

The control law is formulated by the inversion of the above mass matrix. $\ddot{y}_{m_3}$ and $\ddot{q}_a$ can be linearized by

$$\ddot{y}_{m_3} = -k_1(y_{m_3} - y^d_{m_3}) - k_2\dot{y}_{m_3} \tag{5}$$
$$\ddot{q}_a = k_3 x^r_{cm} + k_4 \dot{x}^r_{cm} + k_5 x^a_{cm} + k_6 \dot{x}^a_{cm} \tag{6}$$

where $k_i, i \in 1, 2, .., 6$ are control gains. y_{m_3} is the vertical position of the upper body CoM relative to the wheel axis. x^r_{cm} is the relative CoM of the entire body horizontally to the wheel. x^a_{cm} is the horizontal absolute CoM of the entire body.

PSO

PSO is a stochastic optimization technique inspired by social interaction which does not rely on gradient information but rather estimates the search direction through interactions with neighboring particles. We use PSO to search for the robot trajectory that minimizes the maximum wheel displacement (D_{p-p}) and motor inputs over a broad space of gains and reference as needed by Eq. 5 and 6.

The cost is defined as

$$\text{cost} = D_{p-p} + \int_0^T (R_1\tau_1^2 + R_2\tau_2^2)\, dt \tag{7}$$

where $D_{p-p} = \max(q_a) - \min(q_a)$. R_1 and R_2 are scalar weight.

Implementation

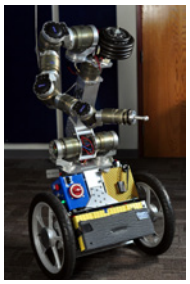
Simulation of PSO including joint position/velocity constraints was done in MATLAB to generate the control parameters and then they were tested on the robot.

Torque cannot be applied directly to the waist motors due to nonlinearity of harmonic drives. Therefore, a PD tracking control with gravity compensation was designed

$$\begin{aligned} \text{Motor input} &= -k_p(q_3 - q_{3_s}) - k_v(\dot{q}_3 - \dot{q}_{3_s}) \\ &\quad + k_g m_3 g r_3 \sin(q_2 + q_3 + d_3) \end{aligned}$$

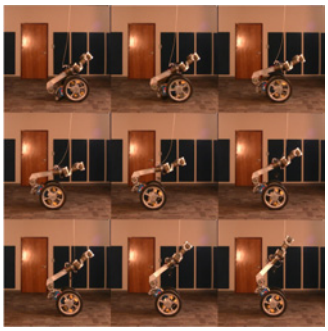
where k_p, k_v, and k_g are control gains. q_{3_s} and $\dot{q}_{3_s}$ are precomputed trajectories of the waist motors obtained from simulation. For the wheel motors, torque command was normally used.

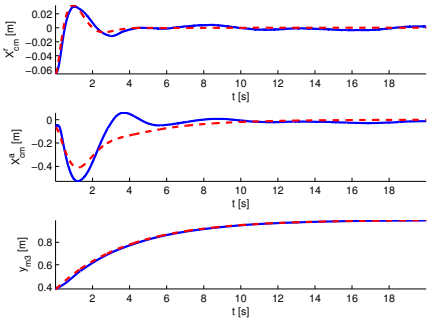
Golem Krang



Standing up with less space

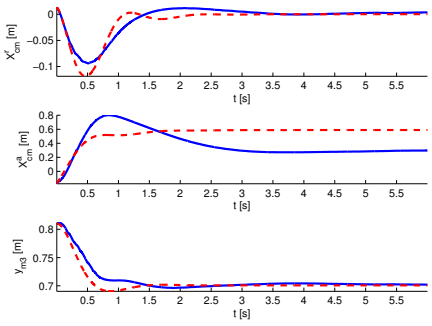
PSO chose k_1 to k_6. The target height was set as $y^d_{m_3} = 1.0$ [m]. The robot required 0.62 [m] to stand up to full height. Note that blue (solid) and red (dash) are of the actual robot and simulation, respectively.





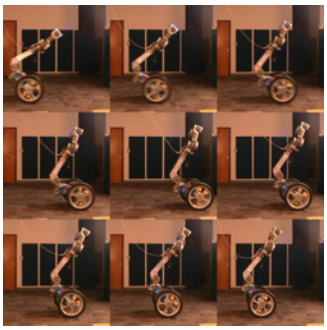
Deceleration by minimizing stopping distance

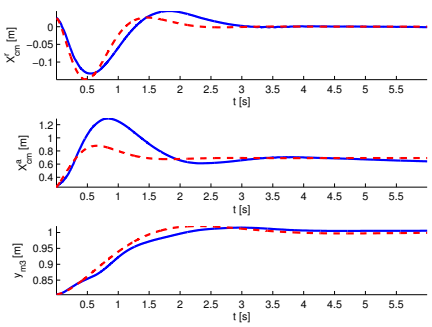
PSO chose k_1 to k_6 as well as $y^d_{m_3}$, except for k_5 which was set to zero. R_2 was also set to zero. The robot lifted its torso down in order to stop within 1.05 [m].



Deceleration by minimizing stopping distance and waist input

PSO chose k_1 to k_6 as well as $y^d_{m_3}$, except for k_5 which was set to zero. The robot lifted its torso up in order to stop within 1.17 [m] and with less torso input.





Advanced Control of Flexible Motion Systems

Intelligent Machine Dynamics Laboratory

The George Woodruff School of Mechanical Engineering

Robotics@GT & Intelligent Machines

IMDL Intelligent Machine Dynamics Laboratory

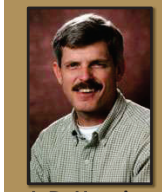
Advisor

Dr. Wayne Book

Student

Brian Post

Research Engineer



J. D. Huggins

Sponsors

Georgia Institute of Technology

CAMotion

National Instruments

Project Goal

To improve the utility and performance of flexible motion systems by enabling high-speed precision manipulation through advanced control methods.

Motivation

Flexibility is often an unavoidable limitation when large-workspace high-speed manipulation is required. Flexible systems are prone to undesirable vibration. To mitigate these effects, they are often operated well below their capable operating speeds. This not only limits the efficiency of these systems, but also hinders their functionality, as dexterity, performance, and workspace size are sacrificed for accuracy and reliability.

Strategic Focus

Thrust 1: Robust Estimation Strategies

• Flexible system models are either too complex for control or fail to capture the true behavior resulting in inaccurate state estimates.

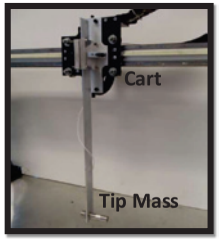
• ***Methods for estimating the state of the system should be robust to disturbances, parameter variation, and modeling inaccuracies.***

Thrust 2: Flatness Based Control (FBC)

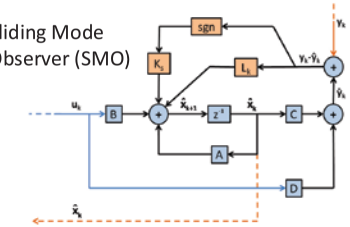
• Feedback control improves vibration reduction and the ability to correct for disturbances. However, feedback compensation requires the existence of vibration before any effort is applied.

•***Potential control structures should utilize command trajectories that result in a reduction of undesired and potentially damaging vibration.***

Robust State Estimation

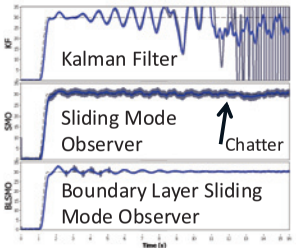


Sliding Mode Observer (SMO)



A robust estimation strategy has been developed. Built on the foundation of a Kalman filter, the estimator employs a sliding mode algorithm to accommodate bounded nonlinearities and changes in system parameters.

Response to 75% Change in Tip Mass



State Update Equations

Kalman Filter (KF)

$$\dot{\hat{x}} = A\hat{x} + Bu + L(y - \hat{y})$$

SMO

$$\dot{\hat{x}} = A\hat{x} + Bu + L(y - \hat{y}) + K_s[\mathrm{sgn}(y - \hat{y})]$$

$$\text{where: } K_s = \rho P^{-1} C^T$$

Boundary Layer SMO

$$\dot{\hat{x}} = A\hat{x} + Bu + L(y - \hat{y}) + S$$

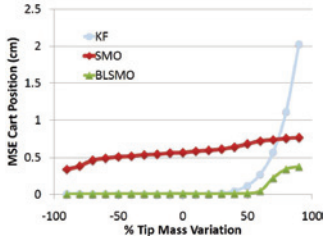
$$\text{where: } S = \begin{cases} \rho P^{-1} C^T \mathrm{sgn}(y - \hat{y}), & \|y - \hat{y}\| > \lambda \\ \rho P^{-1} C^T \dfrac{(y - \hat{y})}{\lambda}, & \|y - \hat{y}\| < \lambda \end{cases}$$

A boundary layer is incorporated to reduce the chatter, below which the discontinuous signum is replaced with a linear slope.

Results:

• The SMO and BLSMO improve robustness relative to the Kalman Filter

• The BLSMO reduces the chatter of the SMO

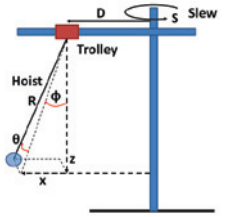


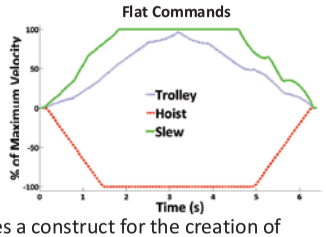
Future Work:

• Expanded experimental observer robustness studies

• Adaptive sliding mode gain adjustment

Flatness-Based Control (FBC)



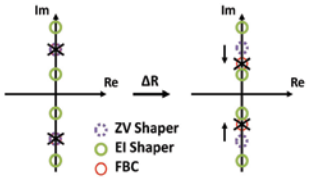


Differential flatness provides a construct for the creation of vibration limiting commands for flexible systems. The selection of a "Flat Output" allows for the specification of the input and state without the solution of a differential equation.

$$x = A(y, \dot{y}, \ddot{y}, \ldots y^{(n)})$$
$$u = B(y, \dot{y}, \ddot{y}, \ldots y^{(n)})$$

Where y is a flat output

By specifying a desired trajectory with no oscillatory behavior, the commands generated will result in no vibration.



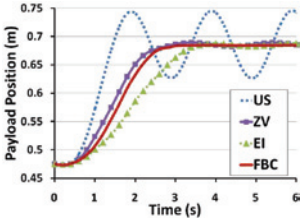
As the system changes, the algorithm tracks the changing system poles resulting in near perfect cancellation.

Results:

• This approach has been applied to a tower crane and benchmarked against standard command shaping techniques

• FBC results in the response speed of a ZV shaper with the vibration rejection of a slower, but more robust, EI shaper.

• Approach requires additional constraints on trajectories



Future Work:

• Expansion of the approach to multiple modes of vibration

• Application to human operated systems

• Refinement of technique for systems with performance limits

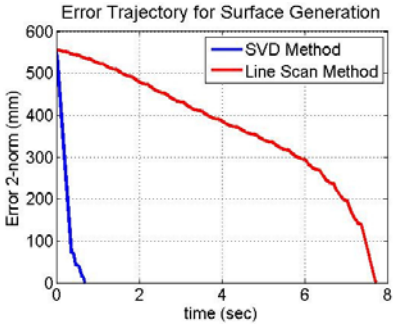
Simultaneous Control of Large Arrays of Actuators with Reduced Control Inputs

Ryder Winck (ryder@gatech.edu) Advisor: Dr. Wayne J. Book (wayne.book@me.gatech.edu)

Background and Motivation

- **Digital Clay is a pin array HMI** that uses a row-column input structure to control the pins.
- The row-column structure allows **control of NxM actuators with N+M inputs** (e.g. 100 actuators controlled using 20 inputs)
- The **current control method** for the row-column structure is the line scanning method similar to one used for LCD screens.
- The line scanning method **is too slow** for many applications.
- **A new method based on Singular Value Decomposition provides a faster and more visually appealing control strategy.**

Current Results

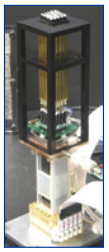


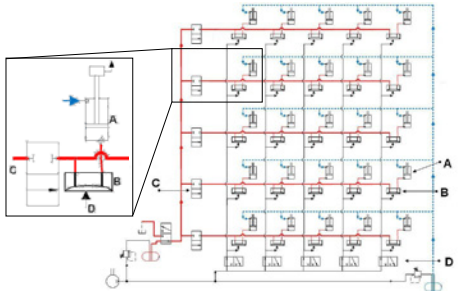
Name	Size	Rank	Line Scan # int. surf	SVD # int. surf	Line Scan time (sec)	SVD time (sec)	% Time Reduction
Identity x50 mm	3x3	3	3	3	1.5	1.5	0
Gradient	3x3	3	3	2	1.8	0.5	70
Parabola	50x50	1	50	1	6.1	0.5	92
Parabola with noise	50x50	50	50	50	7.2	0.7	90
Grid of Squares	50x50	1	50	1	12.0	0.5	96
Grid of Squares Rotated	50x50	40	50	40	24.5	5.9	76
Matlab "peaks"	49x49	4	49	4	15.4	0.7	96
Face	110x82	82	82	82	36.8	2.2	94
Topographical World Map	180x360	180	180	180	68.0	3.7	95

The Singular Value Decomposition Method provides a faster way to control an array of actuators than the line scanning method. In general, the larger the array of actuators the greater the time savings.

5x5 Prototype Overview

- Hydraulic cylinders control pin heights and provide force feedback.
- Can be used as an input and output device.
- Row-Column structure reduces number of valves.

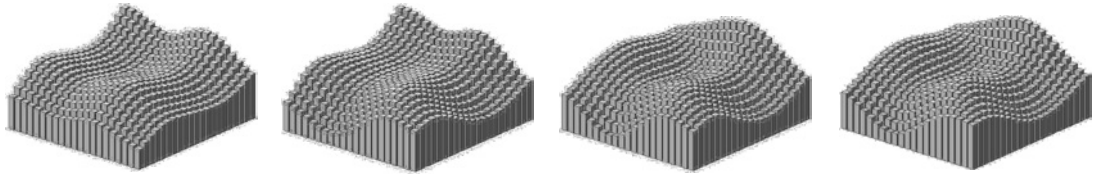




The 25 cylinders (A) each have a control adaptor (B). The fluid to each row of cylinders is controlled using the row valves (C). The control adaptors in each column are open and closed by compressed air control by the column valves (D).

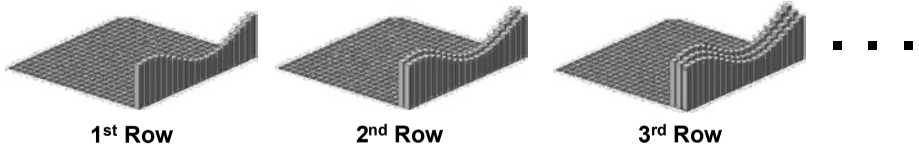
* Zhu, H., 2005, "Practical Structural Design and Control for Digital Clay," Ph.D. thesis, Georgia Institute of Technology, Atlanta, GA.

Singular Value Decomposition (SVD) Method: Generates Final Surface in 0.68 seconds

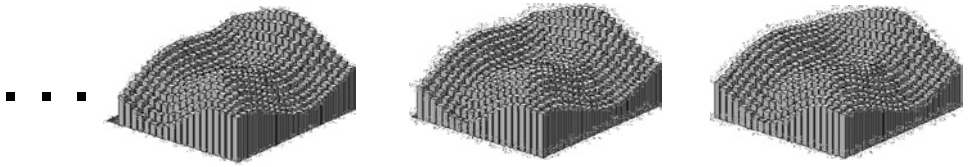


1st Intermediate Surface | 2nd Intermediate Surface | 3rd Intermediate Surface | Final Surface

Line Scanning Method: Generates Final Surface in 7.71 seconds



1st Row | 2nd Row | 3rd Row



19th Row | 20th Row | Final Surface

Future Work

1. Formal written proof that the SVD Method is a time minimum solution.
2. Developed procedure for intermediate surfaces that are outside the physical limits of the actuator.
3. Hardware demonstration of the SVD Method.
4. Reduce surface noise by throwing out the lowest singular values.
5. Rotate the surface to obtain the lowest rank representation.

Special Thanks To:

JD Huggins (Research Engineer, IMDL)

Dr. Haihong Zhu (CTO, Sentrinsic, LLC)

57

CENTER FOR COMPACT AND EFFICIENT FLUID POWER

NSF A National Science Foundation Engineering Research Center

Georgia Institute of Technology | Milwaukee School of Engineering | North Carolina A&T State University | Purdue University | University of Illinois, Urbana-Champaign | University of Minnesota | Vanderbilt University

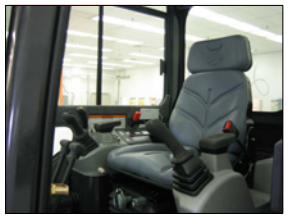
CCEFP
Project 3A.1

Robotics@GT & Intelligent Machines

Multi-Modal Human-Machine Interfaces

Investigators: Dr. Wayne Book, Mark Elton, Georgia Tech

How can multi-degree of freedom hydraulic machines be more effectively controlled? Can effective control increase machine efficiency?



Current Excavator State of the Art

Current state of the art has non-intuitive controls with gentle learning curves that result in:

- Extensive training courses for new operators
- Lengthy experience required to become an expert operator, and experts continue to make mistakes
- Operators are trained on a limited number of machines

More intuitive and easier to use controls will increase the ubiquity of fluid power. Better controls can also increase fuel economy of multi-degree of freedom hydraulic machines.

Even expert operators make errors. More intuitive human-machine interfaces reduce the cognitive load on the operator. Reducing the cognitive load reduces the number of errors. Reducing the number of errors allows the same task to be done in less time and more efficiently. Doing the same task in less in time and more efficiently results in savings in time and money.

Time:

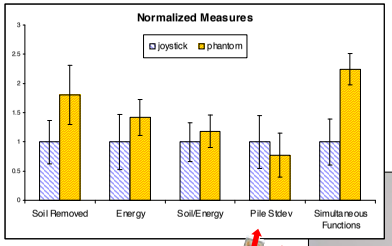
- Current job completed faster
- Jobs waiting on current job completed sooner
- One operator does multiple tasks with different machines

Money:

- Fuel
- Equipment rental
- Equipment maintenance/job
- Operator pay

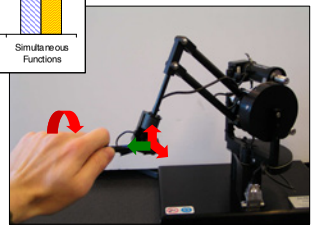
Test Bed 1 at Purdue

Result: 81% time savings and 18% fuel savings!

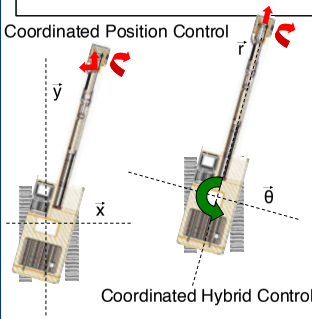


81% more soil removed with 42% more energy consumed results in 18% greater task efficiency

Phantom implements two coordinated control modes



Coordinated Position Control



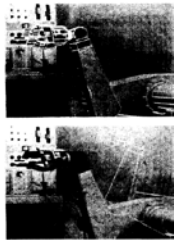
Coordinated Hybrid Control

Haptic Feedback Schemes

- Haptic Wall
- Virtual Spring
- Digging Force Reflection
- Spring Return

Future Directions: Efficient and Effective Supervisory Position Control with Delay

Why position control? Position control has been proven to be more intuitive and effective than rate control, but is not implemented on wide workspace or slow manipulators.

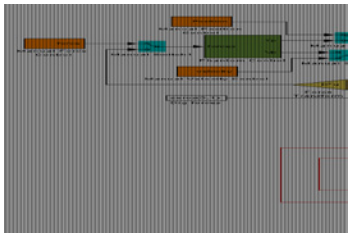


Why supervisory control? Computers are better at crunching numbers and finding optimum paths to maximize fuel economy. Humans are far better at dealing with unstructured environments. Supervisory control plays to the strengths of both.

Why teleoperate? Excavators and other hydraulic equipment are often used in hazardous or remote areas where operators do not want to be. Teleoperation adds time delay to the system.

Operator workstation has been developed for testing new controllers

The simulator includes models of the excavator's mechanical and hydraulic systems, and a new soil-interaction model that better mimics all interactions between the bucket and the soil.



Bobcat

Donations include CAD models and a fully functional mini-excavator.

NCAT collaborates on operator comfort issues.

PURDUE UNIVERSITY

Purdue University is developing TB1 and has supplied modeling measurements.

Robotics@GT & Intelligent Machines

CENTER FOR COMPACT AND EFFICIENT FLUID POWER
A National Science Foundation Engineering Research Center

Georgia Institute of Technology | Milwaukee School of Engineering | North Carolina A&T State University | Purdue University | University of Illinois, Urbana-Champaign | University of Minnesota | Vanderbilt University

Robotics@GT & Intelligent Machines

Modeling and Compensation for Biodynamic Feedthrough in Backhoe Operation

Investigators Heather Humphreys, Prof. Wayne Book and J.D. Huggins, Georgia Institute of Technology

Goal

To develop a backhoe user interface that provides effective and robust compensation for biodynamic feedthrough

Research Questions

How can mobile hydraulic equipment operator interfaces be more effective and efficient?

- *What are the mechanisms through which biodynamic feedthrough affects backhoe operation?*
- *What methods of compensation for biodynamic feedthrough are most effective in backhoe operation, and how can they be applied?*
- *How can haptic feedback and coordinated control be applied to hydraulic machinery to enhance operator capabilities?*

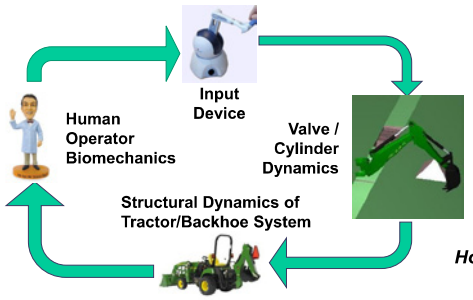
Strategic Alignment and Testbed Demonstration

- **Provides heavy equipment operators with an interface that is easier to use, thus making the equipment operation more effective and efficient**
- **Compensation for biodynamic feedthrough, as well as coordinated control with haptic feedback, can be implemented on the excavator testbed**

Collaborators/Sponsors

NSF · HUSCO INTERNATIONAL · IMDL Intelligent Machine Dynamics Laboratory · John Deere · Georgia Institute of Technology · North Carolina A&T State University

Biodynamic Feedthrough



What is biodynamic feedthrough?

In some operator-controlled machines, motion of the controlled machine/vehicle excites motion of the human operator, which is fed back into the control device, causing unwanted input and sometimes instability; this phenomenon is termed *biodynamic feedthrough*.

How does it relate to fluid power?

Biodynamic feedthrough often causes undesirable oscillations in mobile heavy equipment operation.

Testbed

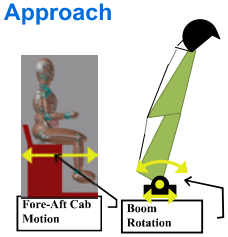
Haptically Enhanced Robotic Excavator (HEnRE)

- **Coordinated position control**
- **Haptic feedback**
- **Low-cost components**

Initial human factors testing revealed significant problems with biodynamic feedthrough, resulting in undesirable oscillations in output.

HEnRE Testbed with Soil Bins

Approach



Single Degree of Freedom Analysis

Fore-aft cab excitation resulting from boom motion with small angle approximations

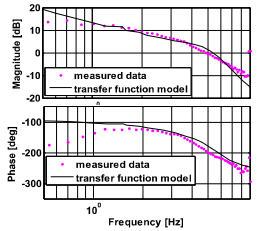
1. ***Create models of the biodynamic feedthrough, represented as transfer functions for each component***
2. ***Test a variety of possible compensation methods, first using simulations***
3. ***Implement the most effective and feasible options in hardware***
4. ***Perform human factors testing to evaluate effectiveness***

Progress

Modeling

- **Determined transfer function models empirically for:**
  - **Valve/cylinder**
  - **Backhoe/tractor structure**
  - **Human operator**
- **Assembled models into a full system dynamic simulation**

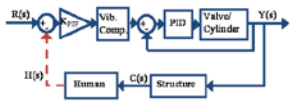
LifeMOD Human Operator Dynamic Simulation



Bode Plots for Valve/Cylinder Model and Measured Data

Compensation Development

- ***Developed 2 types of controllers***
- ***Verified controller designs in simulation using system model with biodynamic feedthrough***
- ***Verified stability for each controller variation***

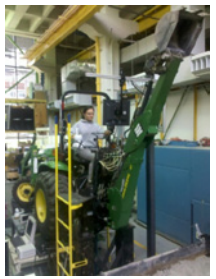


PID Controller with Input Shaper

- Uses input shaping to minimize cab structural excitation
- No additional hardware or sensors needed

Full State Feedback Contoller

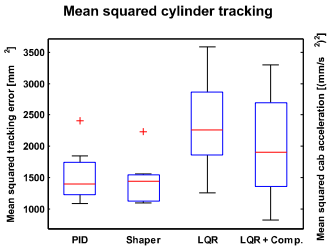
- Optimal LQR control design
- Active cancellation of cab acceleration using accelerometer measurement

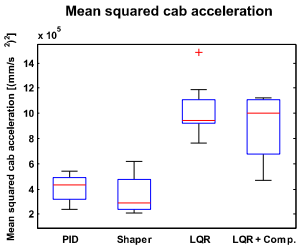


Human Subject Test Setup

Compensation Design Results

- **Tested compensation methods in hardware**
- **Experiment 1: Verified vibration compensation and tracking performance in hardware with software input (no human operator)**
- **Experiment 2: Performed human subject pilot study to verify controller performance in terms of human operator tracking ability**





Next Steps

1. ***Perform larger scale human subject testing, with some modifications to previous experiment***
2. ***Perform additional statistical analyses on human operator study results***
3. ***Investigate ways to expand beyond a single degree of freedom***
4. ***Investigate robustness to parameter variations***

59

CENTER FOR COMPACT AND EFFICIENT FLUID POWER

A National Science Foundation Engineering Research Center

Georgia Institute of Technology | Milwaukee School of Engineering | North Carolina A&T State University | Purdue University | University of Illinois, Urbana-Champaign | University of Minnesota | Vanderbilt University

Robotics@GT & Intelligent Machines

CCEFP Test Bed 4: Compact Rescue Robot

Hannes Daepp, Dr. Wayne Book (Georgia Institute of Technology)
Keith Wait, Dr. Michael Goldfarb (Vanderbilt University)

Test Bed Purpose

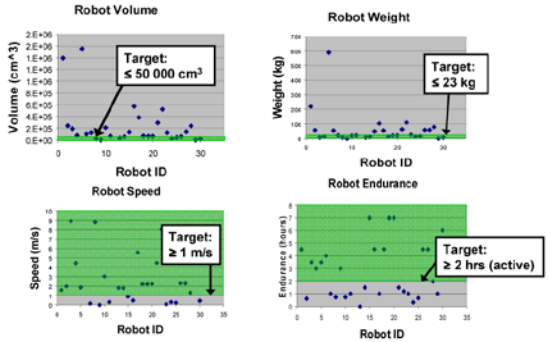
- Demonstrate a self-contained, self-powered, fluid-power-actuated (sub) human-scale legged robot with significantly improved operational characteristics relative to an electrically actuated version.

Why	Where	How
Better performance at the human-scale Provide improved*: • System-level power density • System-level force density • Controllability, operability	• Construction • Forestry • Agriculture • First Responders • Defense • … and many more	• Control aimed at precision, finesse • Techniques for dealing with changing dynamics • Development of interfaces geared towards multi-DoF fluid-powered systems

*relative to an electrically actuated version

Motivation for Fluid Power in Rescue Robotics

- Rescue robotics provides a particularly compelling and comprehensive example of an application that would benefit from fluid power.



- Projected benefits vs. electric are quantitatively apparent:

System	Run Time (hrs)	Mass (kg)	Extra Weight (re. to lightest version)
Electric	**3**	**21**	**10.9**
IC Engine Hydraulic	3	23.1	13.0
HGVP Hydraulic	3	17.7	7.6
FPC	**3**	**10.1**	**0**
Electric	**10**	**36.5**	**24.0**
IC Engine Hydraulic	10	25.9	13.4
HGVP Hydraulic	10	25.2	12.7
FPC	**10**	**12.5**	**0**

Operator Interface

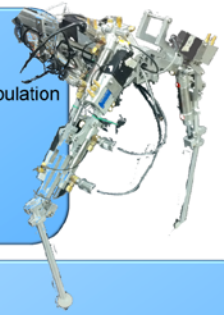
Research
- User control of multi-DoF fluid-power actuated hardware for complex tasks/difficult environments

Design
- Two 3 DoF haptic joysticks coupled with feedback
- High level Human-in-the-Loop control provides technique for human to direct locomotion or manipulation

Two-Legged Robot

Research
- Control of multi-DoF pneumatically actuated systems aimed at precise control and object manipulation

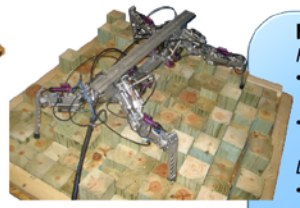
Design
- Originally an interim platform with an alternate design/control strategy from four-legged version
- Two legs, pneumatically actuated
- Pressure and position sensors



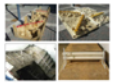
Four-Legged Robot

Research
- Control of multi-DoF pneumatically actuated systems with minimal sensing
- Legged motion

Design
- Four-legged pneumatically-actuated robot capable of walking and manipulation
- Actuation strategy combines cylinder + damper for higher gain margin, reduced sensing need

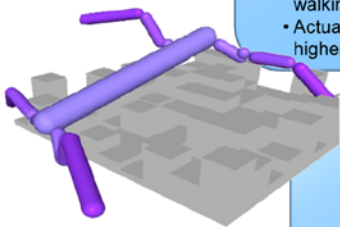


Targets



Design	
Weight	23 Kg
Volume	50000 cm³

Performance	
Speed	1 m/s
Life Per Charge	2 Hours
Maneuverability	Transportable to remote site Able to Negotiate terrain obstacles Capable of accessing rooms through jammed doors
Versatility	Unpin a human dummy trapped by a 230 kg load Accurately position an end of limb as though carrying a sensor Apply controlled light forces in 3 axes of motion



Above: SrLib dynamic simulation
Below: Block diagram of pneumatic actuator model integrated with SrLib

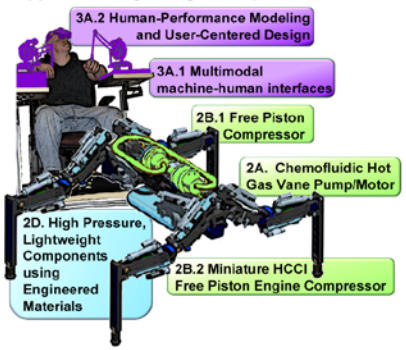
Dynamic Simulation

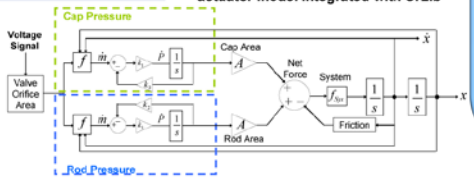
Research
- Modeling and control of multi-DoF pneumatically actuated legged robot

Design
- OpenGL/C++ Simulation Package (SrLib) models dynamics of robot and environment interactions
- Integrates actuator model developed in MATLAB/Simulink to incorporate fluid power dynamics
- Provides a safe, effective, and time-efficient manner to vary design parameters

CCEFP Project Implementations

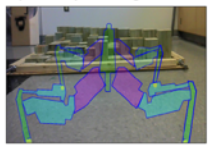
The testbed incorporates CCEFP projects targeting user interfaces, compact power supplies, and lightweight components





Collaborations

1. Human-in-the-Loop (HIL) MPC based controller integrated with a basic gait provides an architecture for guided locomotion (with R. Chipalkatty, M. Egerstedt, GT)
2. "Bird's Eye View" Augmented vision allows user to see robot projected into environment (H. Mizumoto, Kyoto University)



Georgia Institute of Technology
University of Minnesota
Vanderbilt University
FESTO
BIMBA
Kyoto University

60

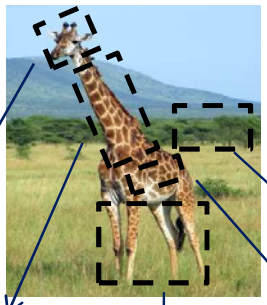
Attribute Learning from Human Interaction

Kaushik Subramanian and Andrea L. Thomaz

Socially Intelligent Machines Lab, Georgia Institute of Technology, GA

Motivation

1. Intuitive to represent objects in terms of their defining attributes.
2. Allows human interaction to take place at a very high-level.
3. Attributes are transferrable across multiple domains.

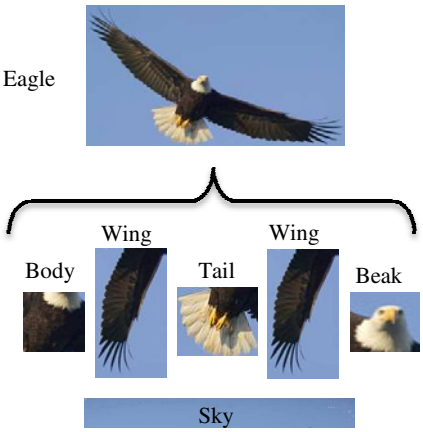


Giraffe = has **horns** and a **long neck** and has **four legs** and is **spotted** and eats **leaves**.

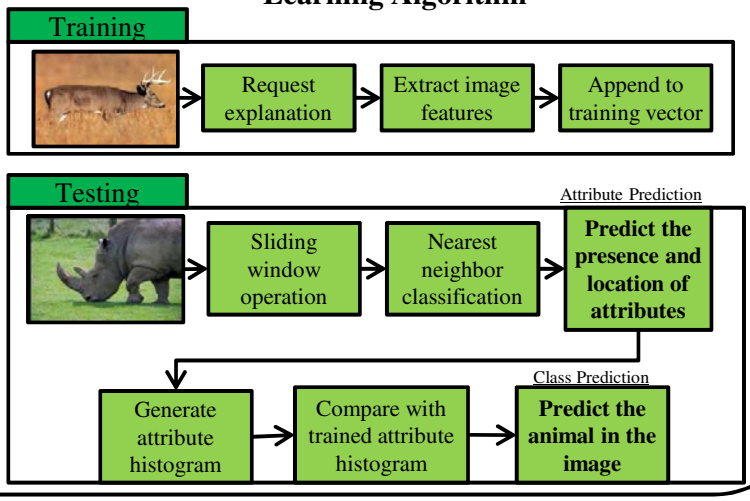
Challenges

1. Learning to predict attributes.
2. Limiting the number of human interactions.
3. To generalize the labeled data to unseen examples.
4. To reuse the learned attributes in other scenarios.

Human Interaction



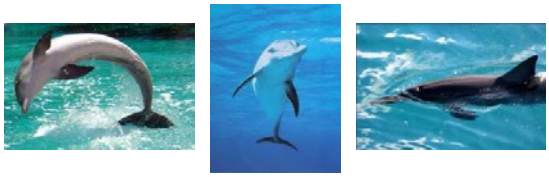
Learning Algorithm



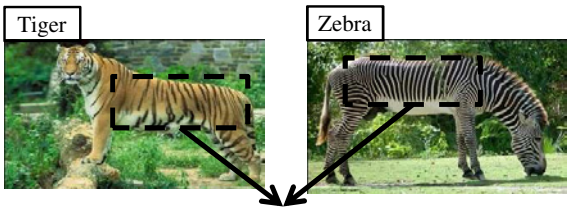
Advantages

Common Attributes = water, fins, flippers, tail, grey

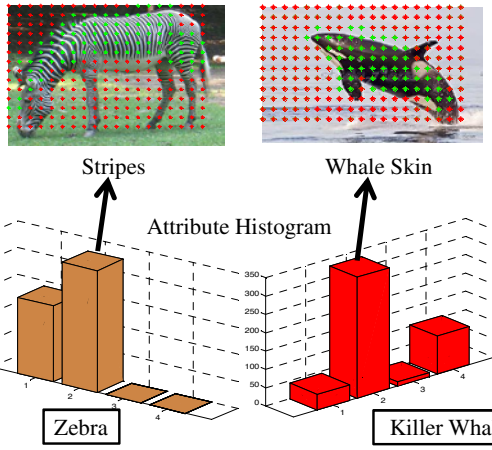
Generalization over different images



Transfer of learned attributes



Results



This work has been sponsored by a Honda 2011 Initiation Grant

Trajectories and Keyframes for Kinesthetic Teaching: A Human-Robot Interaction Perspective

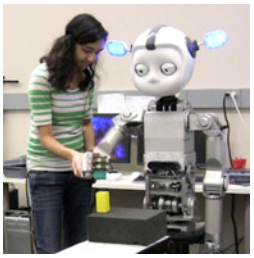
Jae Wook Yoo, Maya Cakmak, Baris Akgun and Andrea L. Thomaz

Socially Intelligent Machines Lab, Georgia Institute of Technology

Motivations & Goal

Robot Learning from Demonstration (LfD)
- Personal robots won't be possible to be preprogramed for every specific task
- Every day users should be able to teach robots without programming expertise

Kinesthetic Teaching
- physically guiding robots to demonstrate a skill
- avoiding correspondence problem
- being restricted to robot's capabilities

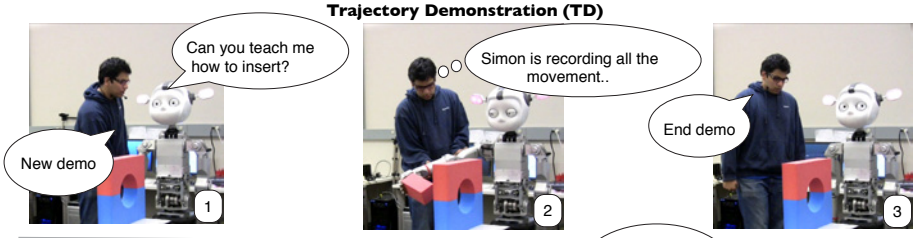


Challenges in Kinesthetic teaching
- Everyday users do not have experience manipulating a robot arm with many degrees of freedom.
- The usability of the interaction interface has not been explored in depth

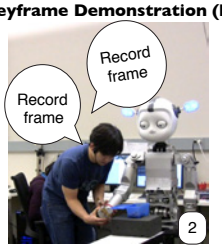
Goal of the Study
- Identifying challenges that everyday users are faced with common methods in the field
- Proposing improvements to these methods to increase their usability.

Implementation & Experiments

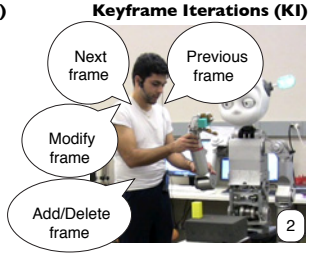
Trajectory Demonstration (TD)



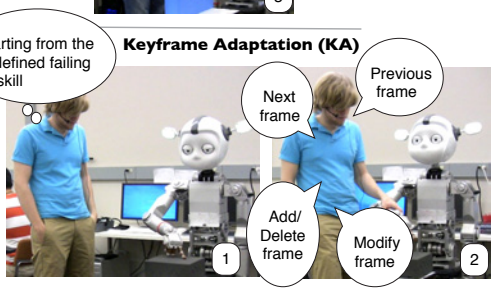
Keyframe Demonstration (KD)



Keyframe Iterations (KI)



Keyframe Adaptation (KA)

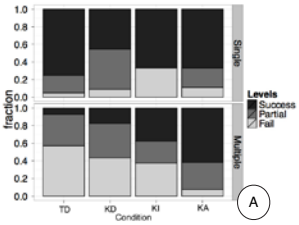


Goal-oriented skills
- Insert, Close, Stack, Touch

Means-oriented skills
- Salute, Throw, Beckon, Raise-hand

Results

Fraction of participants who achieved different levels of success

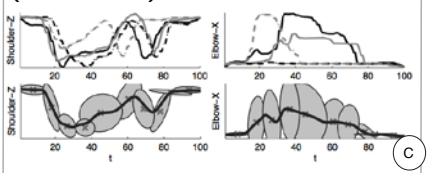


Goal- and Means- skills

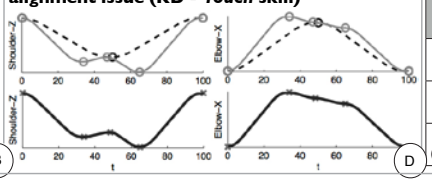
	Goal-. skills	Means-. skills	t-test
Avg. # of demo	2.37 SD=1.45	1.22 SD=0.53	t(84)=6.18 p < 0.001
Avg. distance between keyframes	0.12 (47% smaller)	0.27	t(38)=-3.94 p < 0.001

B

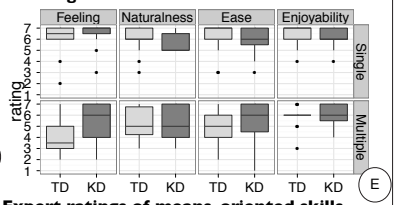
Examples of the temporal alignment issue (TD - *Close* skill)



Examples of forgetting obstacle avoidance, and of good handles of the temporal alignment issue (KD - *Touch* skill)



Subjective ratings of TD and KD for goal-oriented skills



Expert ratings of means-oriented skills (Median and Coefficient of dispersion)

Cond.	Expert	*Appropriate Emphasis*	*Communication of Intent*	*Closeness to Perfection*
TD	1 2	5.5 (0.27) 3 (0.29)	5 (0.33) 3.5 (0.38)	5 (0.35) 4 (0.3)
KD	1 2	6 (0.21) 4(0.21)	6 (0.17) 4 (0.24)	6 (0.2) 5 (0.22)
TD vs. KD (Wilcox s.r. test)		Z=2.679 p=0.10	Z=2.677 p=0.11	Z=2.796 p=0.03

F

Advantages of keyframe-based LfD
- TDs were preferable for goal-oriented skills when the skill is taught with a single demonstration (A)
- Multiple TDs can result in alignment problem (A) (C)
- KDs were good at handling multiple demonstrations (A) (D)
- KDs can let the user make temporally local modifications on a skill (KI) (A)
- People were able to use the KI starting from a skill that was not their own (KA) (A)
- KDs may result in preferable means-oriented skills (F)
- Participants like both TD and KD (E)

Comparison of Goal- and Means- skills
- Different objective functions are used for each skill type (B)
- Characteristics of provided keyframes are different for each skill type (B)

Future work
- **Timing in Keyframe-based LfD:** timing, and speed of movements.
- **Different types of keyframes**: goal-keyframes and waypoint keyframe.

**This study is submitted to the *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2011*

Robotics@GT & Intelligent Machines

Social Exploration Strategies for Robots Learning from Human Partners

Karl Jiang and Andrea L. Thomaz

Socially Intelligent Machines Lab, Georgia Institute of Technology

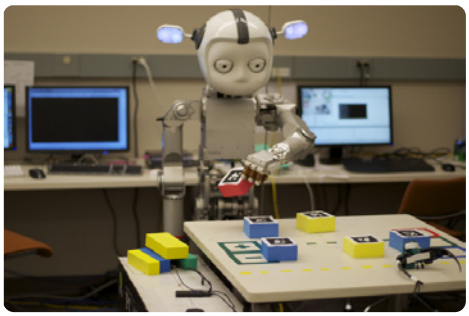
Motivation

Object Affordances

- Robots cannot be preprogrammed for every possible environment
- Must dynamically learn effects of actions on objects

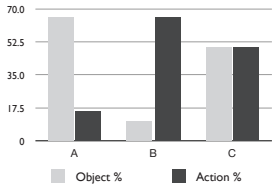
Social Exploration Strategies

- Humans use different strategies to learn affordances from their peers
  - *Stimulus Enhancement / Emulation:* Pay attention to the objects peers act on and the effects of their actions
  - *Mimicry:* Pay attention to actions peers use
  - *Imitation:* Imitate both the action used and the object it is used on
- Previous study [1] showed different exploration strategies are effective in different environments, with robot teacher
- Are the same strategies useful in interactions with humans?



Implementation & Experiments

Experimental Environments



Object % reflects percentage of objects capable of producing an effect. Action % is the percentage of effect-producing actions. Different-goal means the human and robot were trying to achieve different effects.

Experiments

- Different environments modeled by different concentrations of objects/actions that produce effects
- Strategies chosen based on what is expected to do well in each environment

Implementation

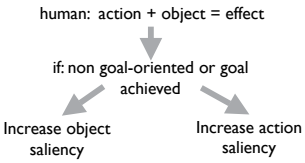
- Objects are augmented reality toys
- Actions are AR tool blocks which act on objects
- Operational space control constraints to grab blocks

Conditions Run

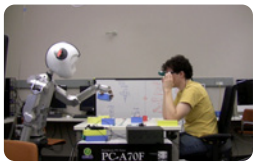
Environment	Strategy 1	Strategy 2
A	Stimulus Enhancement	Mimicry
B	Stimulus Enhancement	Mimicry
C (different-goal)	Goal-oriented Imitation	Non goal-oriented imitation

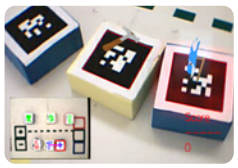
Hypotheses

- Stimulus Enhancement yields better recall/accuracy in environment A
- Mimicry over S. Enhancement, in environment B
- Goal-oriented over non goal-oriented, in C



Strategy	Action chosen by	Object chosen by
S. Enhancement	Random choice	Object saliency
Mimicry	Action saliency	Random choice
Imitation	Action saliency	Object saliency





Preliminary Observations

Subject Confidence in Simon

'How well do you think Simon learned about the objects?'

- Strong correlation between how well subjects reported they understood the object affordances and how well they thought Simon did
- Currently analyzing: correlation with how well Simon actually learned the affordances

Observed Teaching Behavior not Picked Up by Exploration Strategies

- Subject places block on Simon's side (but does not act on an object)
- Subject emphasizes a certain property of an object
- Subject hides an irrelevant object

Reference: [1] M. Cakmak and A.L. Thomaz, "Exploiting Social Partners in Robot Learning." Autonomous Robots, 2010.

Robotics@GT & Intelligent Machines

Mixed-initiative Learning for Robots

Maya Cakmak and Andrea L. Thomaz
Socially Intelligent Machines Lab, School of Interactive Computing, Georgia Institute of Technology

Motivation

Robots that learn new things from everyday people.

Example "Clean up the kitchen" task

- Different in every home
- Only users know what they want

Challenges

Human
- Limited time
- Limited patience

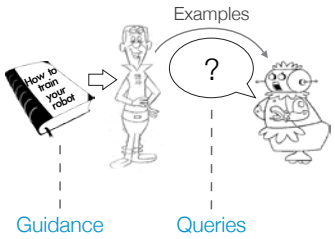
Robot
- Limited perception
- Limited model of human

***Goal**: Learn as much as possible from as few examples as possible!*

Approach

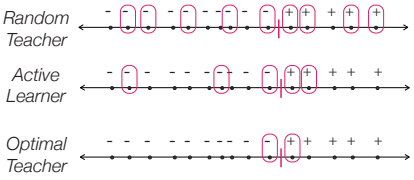
Share responsibility of choosing examples between human and robot, combine strengths of both.

Mechanisms



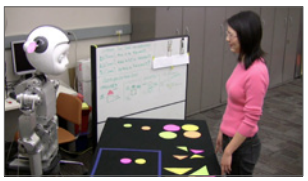
Machine Learning

- Active learning
- Optimal teaching



Human-Robot Interaction

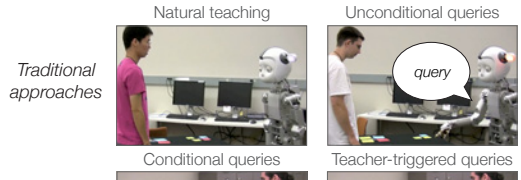
- Human teaching behavior
- Subjective preferences



Findings

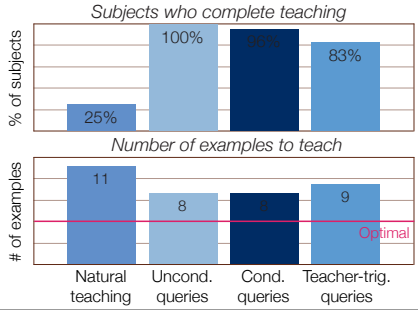
Human teaching can get close to optimal with proper usage of queries and guidance.

Experiment 1

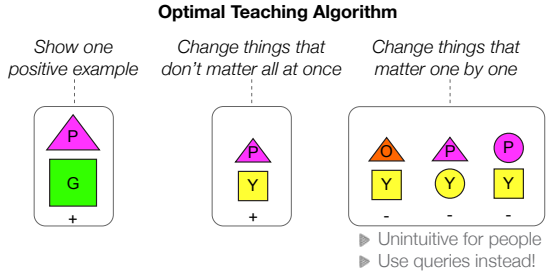


- Balanced
- Less effort

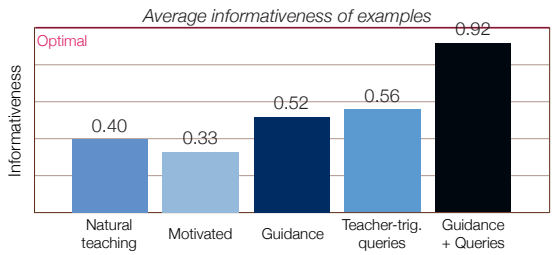
- Increased control over teaching



Experiment 2



- Unintuitive for people
- Use queries instead!



References

[1] M. Cakmak and A.L. Thomaz "Optimality of Human Teachers for Robot Learners" International Conference on Development and Learning, 2010.
[2] M. Cakmak, C. Chao and A.L. Thomaz "Designing Interactions for Robot Active Learners" IEEE Transactions on Autonomous Mental Development, Vol. 2, No. 2, June 2010.

Human-like Motion For Anthropomorphic Robots

Michael J. Gielniak, C. Karen Liu,
and Andrea L. Thomaz

Georgia Institute of Technology
School of Interactive Computing

Robotics@GT & Intelligent Machines

Motivation

Robots interacting with humans can benefit from communicating in a manner that is socially relevant and familiar to their human partners. Believable behavior establishes appropriate social expectations. Our work addresses the overall problem of how to generate believable or human-like motion for an anthropomorphic robot.

Research Problem

Anthropomorphic robot bodies are different from human bodies. The challenge is to create human-like motion for robots autonomously from minimal input information, despite the differences between robots and humans. Our work focuses on a series of algorithms that work toward improving motion quality for social robots.

Approach

We divide the problem of human-like robot motion for anthropomorphic robots into a series of discrete components, each of which present unique challenges.

(1) Social robots must use motion as a communication channel to communicate with human partners

(2) Motion must exhibit variety, as opposed to redundancy, so that gestures are never performed the same way twice.

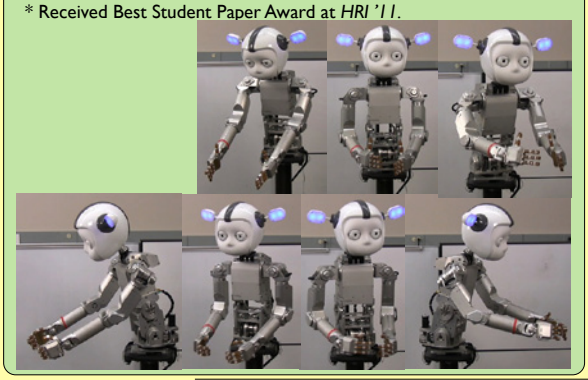
(3) Motion should appear intuitive, recognizable, and familiar to human partners.

(4) Social robots must maintain world constraints in unison with all social constraints.

The Components of Human-like Motion

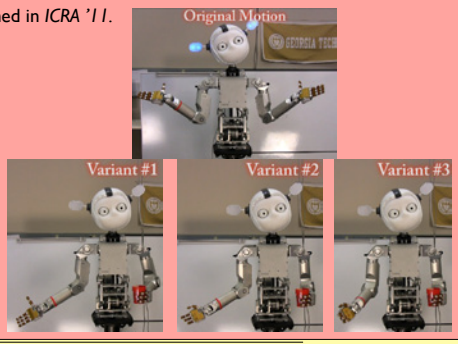
Optimized To Be Spatially & Temporally Coordinated

* Humans can more accurately label, recognize, and imitate STC-optimized motion than uncoordinated motion.
* Coordinating robot motion with respect to the STC metric makes the motion more human-like.
* Coordinating motion increases recognition accuracy and the ability to imitate.
* Received Best Student Paper Award at *HRI '11*.



Variance in the Presence of Constraints

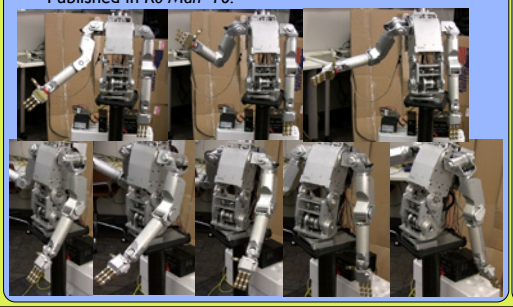
* Optimal control can be used to find a local solution for state-space directions to inject torque-space noise.
* The injection of this noise produces visible differences to motion, thereby creating motion variants similar to the original motion.
* When coupled to operational space control, the resultant motions respect task-based world constraints.
* Published in *ICRA '11*.



Adding Communication

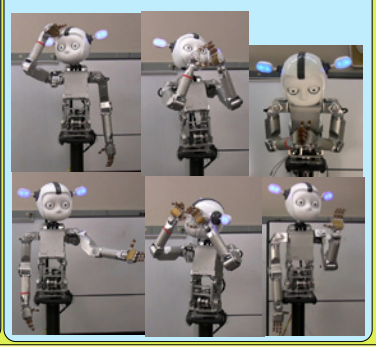
Secondary Action

* Secondary action communicates internal (robot) and external (world) state information such as motor temperature (fatigue), friction, mass, and passivity.
* Eigenanalysis on motion torques allows decomposition of coordinates into highly-actuated and nearly-unactuated.
* Our algorithm projects the highly-actuated torque response onto the nearly-unactuated coordinates so that internal and external force responses are consistent with respect to motion across the entire robot body.
* Published in *Ro-Man '10*.



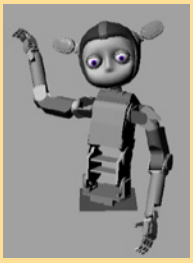
Anticipation

* Motion used for communication contains what we term as the 'symbol,' which is one frame that yields highest motion recognition accuracy.
* By extracting and moving a symbol earlier in motion, humans can consistently label motion intent sooner, which has social interaction benefits such as decreased frustration and increased cooperative task response time from the human partner.



Exaggeration

* Exaggeration is used to draw attention, when something will be communicated via motion.
* Exaggeration moves parts of the body through larger regions of travel in Cartesian space, while minimizing the parts of the body to which attention should not be drawn.
* By "extremizing" coordinates based on actuation spectrum a single parameter can be used to control the amount of exaggeration to induce into a particular motion.



Human-Like Action Segmentation for Option Learning

Jaeeun Shim and Andrea L. Thomaz
Socially Intelligent Machines Lab
Georgia Institute of Technology

Research Statement

Robot learning interactively with a human partner has several open questions, one of which is increasing the efficiency of learning. One approach to this problem in the Reinforcement Learning domain is to use options ,temporally extended actions, instead of primitive actions. We aim to develop a robot system that can discriminate meaningful options from observations of human use of low-level primitive actions, inspired by psychological findings about human action parsing.

Expectation

We expect that the robot agent can have the following two benefits from options detected based on human-like action segmentation mechanism.

1. Speeding up the learning process
2. Solving the sub-problem faster with the better performance

Approach

Motivation from Human Psychology

Humans attend to statistical regularities in small-scale primitive actions to determine action boundary choices.

[Baldwin et al., 2008]

Options and SMDP

MDP (Markov Decision Process) : defined by the primitive actions

Options : temporally extended actions

SMDP (Semi-Markov Decision Process) : defined over the MDP using options

Previous Study

Experimental Design

System Domain
- a. 7 colored objects
- b. 17 primitive actions
- c. 36 discrete states

Evaluation Steps

1. Action Segmentation Session
   - a. robot observes the demonstration of real human subjects and discriminates options (18 participants)
2. Learning Session
   - a. we evaluate whether the detected options are efficient for robot LbD

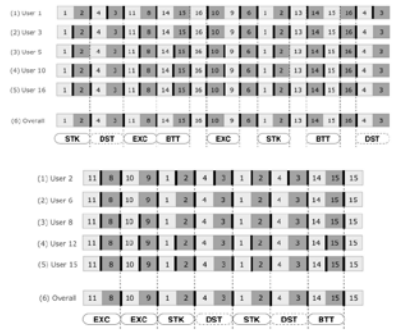
7 objects & Initial State

Goal State for learning

Human-Like Action Segmentation to Find Options

1. The robot identifies options using the human-like action parsing mechanism
2. Action Segmentation
   - a. Generate statistical model of the transition probabilities T_{ij} between the primitive actions during the human demonstration
   - b. Determine the threshold $threshold = \underset{i \in PrimitiveActions}{average} MaxProb_i$
   - c. Detect the action boundaries from a sequence
     $$isBoundary(i,j) = \begin{cases} true & \text{if } T_{ij} < threshold \\ false & \text{otherwise} \end{cases}$$
3. After the action segmentation process, the discriminated meaningful set of primitive actions $\{a_1, a_2, a_3\}$ defines options
   - a. initial state : primitive action a_1
   - b. termination state : primitive action a_3
   - c. policy : a set of primitive actions $\{a_1, a_2, a_3\}$

Action Segmentation Results



Individual data : can't segment the meaningful actions well

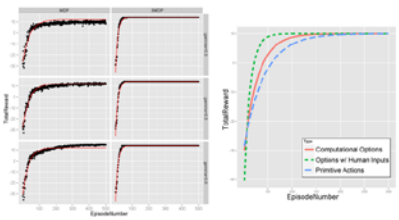
Entire data : segment all the meaningful actions successfully

Simulation for Learning Session

Learning efficiency : how fast robot finds the optimal policy from an initial state to the goal state with the Q-learning algorithm

1. Compare MDP (primitive actions) with SMDP (options)
2. Compare our options from the human-like action segmentation with the computational options from the bottleneck algorithm

Learning Results



Primitive Actions Vs. Our Options

Learning with our options converges to the optimal policy faster than other two.

Our options lead to the better learning efficiency.

*) This research submitted to Ro-Man 2011 and currently under review.

Conclusion and Current Study

From the previous study

1. Robot can find options
   - a. with the human-like action segmentation mechanism
   - b. from the low-level primitive actions
2. Robot can identify options from the action space, not the state space
3. Our options can speed up the learning

Current Study - Assembly Furniture Task

1. We aim that robot can learn the goal-oriented task from the low-level primitive actions
2. Assembly Furniture Task is a good goal-oriented task including many sub-tasks
3. Using the human-like action segmentation, we expect that robot can also effectively solve the sub-problem during the assembly furniture task

Contingency Detection for HRI

Jinhan Lee
Advisors: Aaron F. Bobick* and Andrea L. Thomaz**
Computational Perception Lab* and Socially Intelligent Machines Lab**
Georgia Institute of Technology

Robotics@GT & Intelligent Machines

Motivations & Goal

Contingent Behavior

- Psychological Definition: A person's behavior in response to the behavior of another
- Our definition for HRI: A robot generates an interactive signal to a human partner and detects their response.
- We use a contingency mechanism as an augmented perceptual cue for two interaction scenarios: **Engagement Detection** and **Turn-Taking**

Engagement Detection

- Determining whether a person is interested in having an interaction with a robot

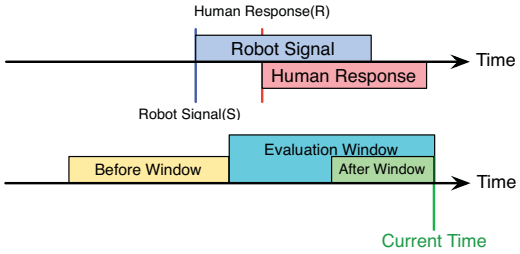
Turn-Taking

- During interaction, determining whether a robot should hold, yield, or negotiate turns at each point in time while performing collaborative or competitive tasks

Engagement Detection

Turn Taking

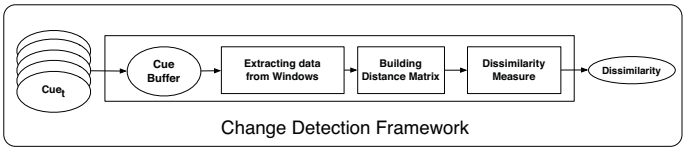
Approach



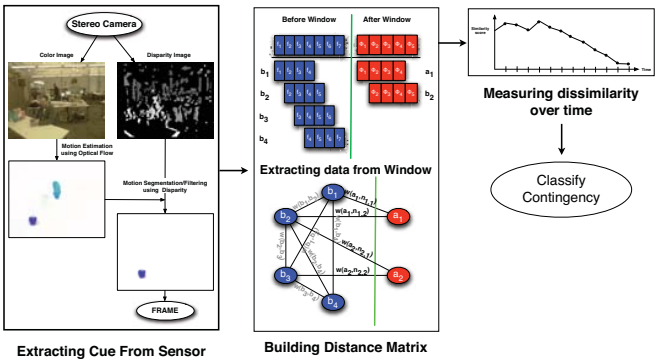
Contingency Detection: determine the presence/absence of a human response (R)

Two strategies: Event Detection and Change Detection

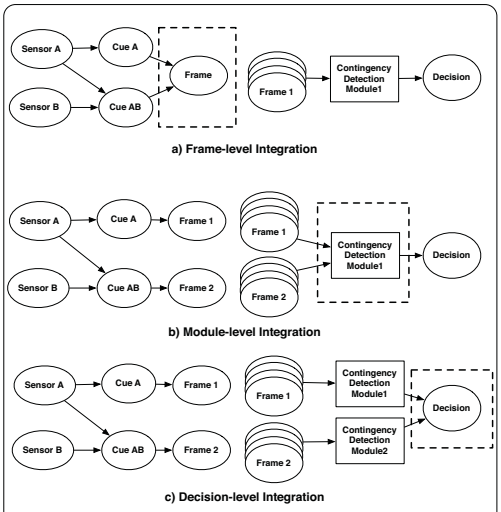
- **Event Detection**: Looking for specific events as a response
  - Analyzing sensor observations within After Window
  - E.g. audio cue: voice on during some fixed window, touch cue: contact sensed during some fixed window.
- **Change Detection**: Looking for behavioral changes before and after the robot's signal
  - Measuring statistical difference between sensor observations in Before Window and After Window.



Change Detection Framework



Our Implementation of Change Detection



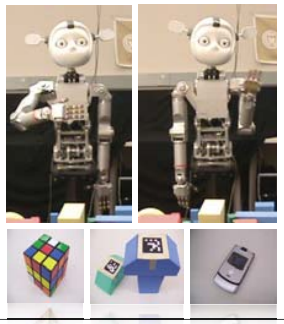
Cue Integration at three different levels using Naïve Bayes Probabilistic Model

Results

Experiments for Engagement Detection+

Setup:
A human is not engaged with a robot while he is doing an assigned task (either playing with a Rubik's cube, playing with foam blocks, talking on cell phone) **The robot** does either waving or beckoning to the human and detects the presence of any responses.

Among 41 interaction sessions from 5 subjects, our classifier detects responses with an accuracy of 80%.



Experiments for Turn-Taking++

Setup:
A human and **a robot** plays an imitation game based on the traditional children's game "Simon-says". During the game phase, the leader says sentences of the structure, "Simon says, [perform an action]" or "[Perform action]".

Among 167 interaction sessions from 11 subjects, our classifier detects responses with accuracy of 85% when depth and motion cues are used together.

+published in *International Conference on Human Robot Interaction (HRI) 2011*
++submitted to CVPR Workshop for Human Communicative Behavior Analysis(CVPR4HB) 2011

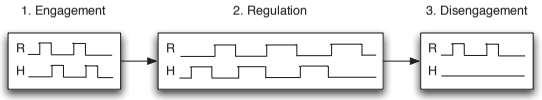
Turn-Taking for Human-Robot Interaction

Crystal Chao and Andrea L. Thomaz
Socially Intelligent Machines Lab, Georgia Institute of Technology

Robotics@GT & Intelligent Machines

Turn-Taking

An interaction can be viewed as being in one of three stages: engagement, regulation, and disengagement.[1]



A contingency detector can tell a robot whether to engage or disengage a person (stages 1 and 3). After engaging, the dyad performs turn-taking to regulate interaction timing (stage 2).

The turn-taking process is highly multimodal and uses many channels for communicating turn state, including gaze, prosody, and whole-body gesture.[2,3,4]

Goals

Interaction improvement
A model for turn-taking may allow robots to have smoother, more efficient interactions that are less frustrating for humans.

Some common error patterns in our interactions include the human waiting for too long, or the human repeating himself:[5]

Ease of programming
For each new domain or task, the programmer should be able to implement only domain-specific behaviors without spending time coding turn-taking behaviors in the instrumental FSM.

Domains

Turn-taking is often investigated in the free conversation domain.[3,4,6] Other interaction domains may have significant effects on dynamics and observation distributions, such as for gaze and motion.

Some domains we are considering include learning from human demonstrations, imitation games, play-based interactions with children, and providing information to passersby.

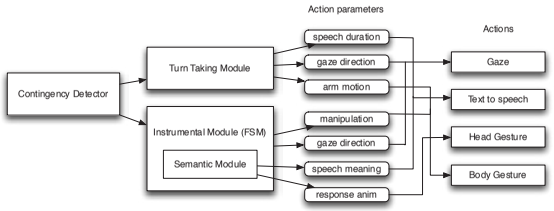
Architecture

We hypothesize that there is some context-free component of turn-taking, which allows us to modularize the behavior and reuse it in a way that is generic to multiple types of interactions.

Components
Contingency Detector – triggers stage 2 modules
Turn-Taking Module – context-free turn regulation
Instrumental Module – task-based finite state machine
Semantic Module – library of ways to communicate a meaning

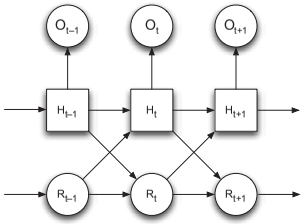
Action Selection
The context-based (instrumental) and context-free (turn-taking) modules generate action parameters, which combine to instantiate actions that are executable by the robot.

For example, speech duration and speech meaning can combine to select a context-appropriate utterance that suppresses the auditor's turn for some time.



Temporal Inference Model

We previously modeled turn-taking as a partially observable Markov decision process. Because of the nature of the horizon, it was difficult to derive good policies. Currently we describe turn dynamics as the following first-order Markov process:



At time *t*:

O_t = observation
H_t = human turn state
R_t = robot turn state

The hidden state H_t is filtered over time.

States: R_t, $H_t \in$ {Seizing, Passing, Holding, Listening}

Observations: Vector of binary features, e.g. O_t = <1,0,1,1,0>
Features: <mic level, gaze direction, speech recognized by grammar, hand motion, prosody derivative>

Conditional probability distributions
Three functions to train from video-coded experiment data, in which the robot is tele-operated to do good/bad turn-taking:

Observation function $P(O_t|H_t)$
how well the sensor data reflects the human turn state H_t
Human transition function $P(H_t|H_{t-1},R_{t-1})$
how the human takes turns
Robot transition function $P(R_t|R_{t-1},H_{t-1})$
how the robot should do turn-taking. a good and a bad model

Extensions

Semi-Markov process: A Markov chain assumes that the times between events are geometrically distributed. Other distributions require an additional node Q_t representing time spent in a state.

Switching model: The human transition function may differ across people; certain people may be more aggressive or more passive turn-takers. The robot can select the maximum likelihood model.

State space: Backchannels could be explicitly represented.

[1] Tronick, E.; Als, H.; and Adamson, L. 1979. Structure of early face-to-face communicative interactions. *Before Speech: The Beginning of Interpersonal Communication*. [2] Cassell, J., and Thorisson, K.R. 1999. The power of a nod and a glance: envelope vs. emotional feedback in animated conversational agents. *Applied Artificial Intelligence* 13:519–538. [3] Mutlu, B. 2009. Designing gaze behavior for humanlike robots. PhD thesis. Carnegie Mellon University. [4] Orestrom, B. 1983. *Turn-Taking in English Conversation*. [5] Chao, C., and Thomaz, A.L. 2010. Turn-taking for human-robot interaction. *AAAI Fall Symposium on Dialog with Robots*. [6] Sacks, H.; Schegloff, E.A. and Jefferson, G. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50:696–735.

Control of Underwater Gliders for Study of Wintertime Phytoplankton Bloom in Long Bay, SC

D. Chang, K Szwaykowska & F. Zhang
School of Electrical and Computer Engineering
Georgia Institute of Technology

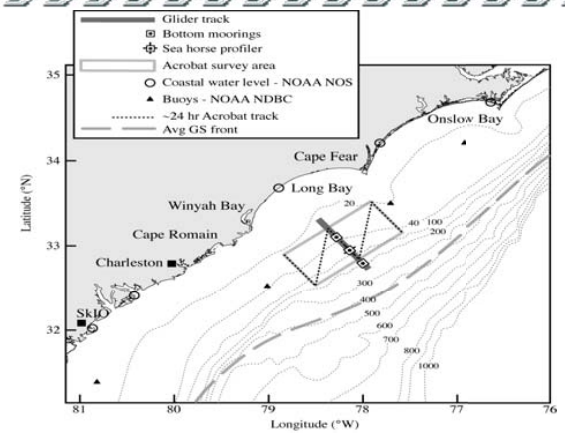
This project is being conducted in collaboration with oceanographers at the Skidaway Institute of Oceanography in Savannah, GA.

Scientific goals:

- Investigation of nutrient sources which support unexpected annual winter phytoplankton blooms off the coast of Long Bay, SC.

Engineering goals:

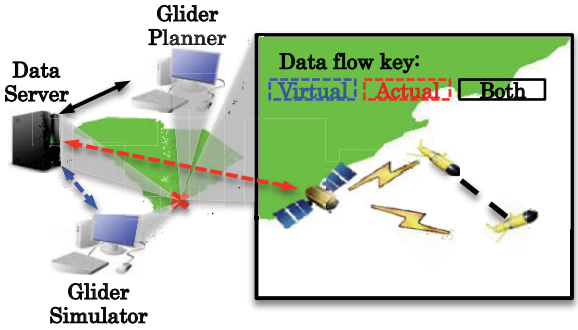
- Autonomous control of two underwater gliders, one being a virtual mooring; the other traversing onshore and offshore, for collection of oceanographic data along specified tracks.
- Use of Ocean General Circulation Models (OGCMs) to simulate glider motion through time-varying and unstructured ocean flow fields.

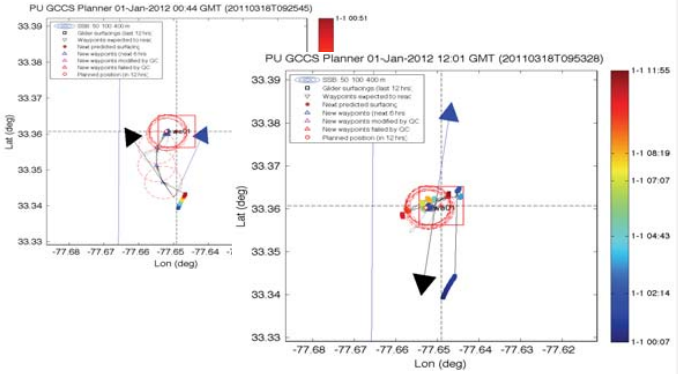


Survey area near Long Bay, SC. The experiment will be conducted in Jan-Feb. 2012.

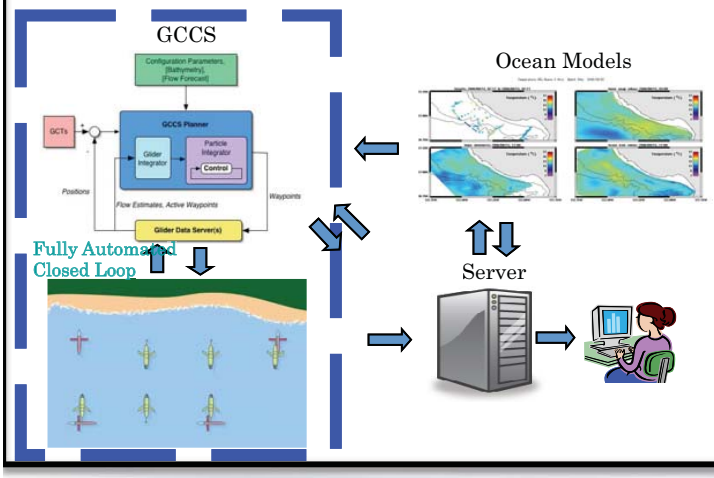
Glider Control:

- The Glider Coordinated Control System (GCCS) is used as a simulation and control platform for underwater vehicles.
- Glider trajectories are simulated using a desired control law and flow data from ocean models. Waypoints are produced based on the generated trajectory and transmitted to the AUV onboard controller via satellite.





Sample AUV trajectory and waypoints simulated using the GCCS, for AUV performing a station-keeping mission ("virtual mooring").

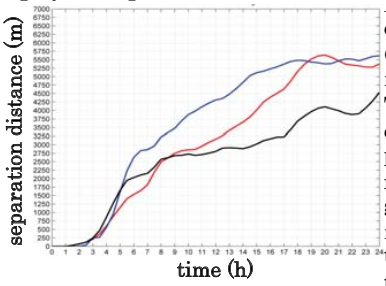


Limits of using ocean models for motion control:

- Errors in modeled flow lead to inaccurate predictions of glider motion.
- We distinguish 2 flow components: mean flow F, and turbulent, small-scale flow v. Expected separation (e) between modeled and physical AUV grows as:

$$\frac{dE[e]}{dt} = E[f(x,t)]$$

(f is error in modeled mean flow F, and x is AUV's physical position).



Data from 2006 ASAP field experiment in Monterey Bay, CA, showing separation between modelled and physical gliders. The separation initially grows exponentially, then slows to polynomial growth after reaching approx. twice the gridsize used by the ocean model (~2 km). We have derived theoretical results to explain this behaviour.

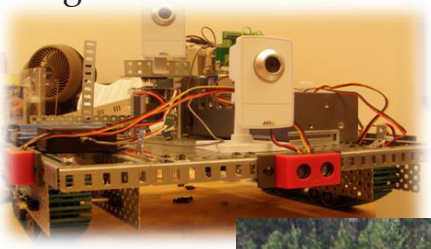
Victoria -Autonomous Surface Vehicle

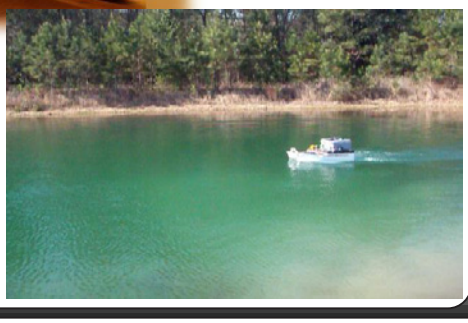
Brandon Groff Dr. Fumin Zhang

Georgia Tech Robotics & Intelligent Machines

Georgia Tech Savannah Robotics has been developing *Victoria* since January 2010. It began as a joint senior design project between electrical, computer, and mechanical engineers. *Victoria* was entered in the 2010 Autonomous Surface Vehicle (ASV) Competition sponsored by AUVSI and ONR.

The control system for *Victoria* was developed in parallel with the hull fabrication. In order to reduce development time, a simple land-based vehicle was used to test hardware and navigation controllers.



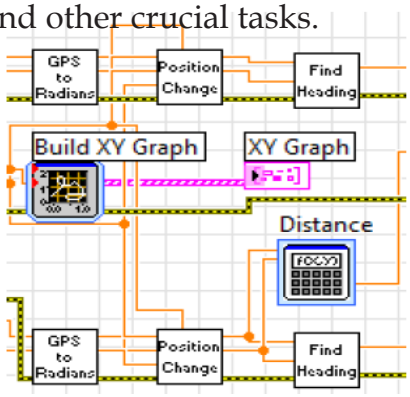


Ongoing Research

- Surveying of inland waterways in Louisiana for residual oil
- Field-tracking
- Mapping
- Hardware in loop simulation
- More advanced navigation controllers
- Cooperative exploration
- Vehicle, thruster & battery modeling (for simulation and improved controller design)

Software Control System

LabVIEW 2010 is used for all software development. A state machine architecture controls the overall tasks *Victoria* performs, while various processes handle sensor readings, data logging and other crucial tasks.



Major Components:

- CompactRIO Embedded Computer & FPGA
- Garmin GPS
- Microstrain IMU
- Sick LIDAR
- Axis Network Cameras
- Crust Crawler Thrusters

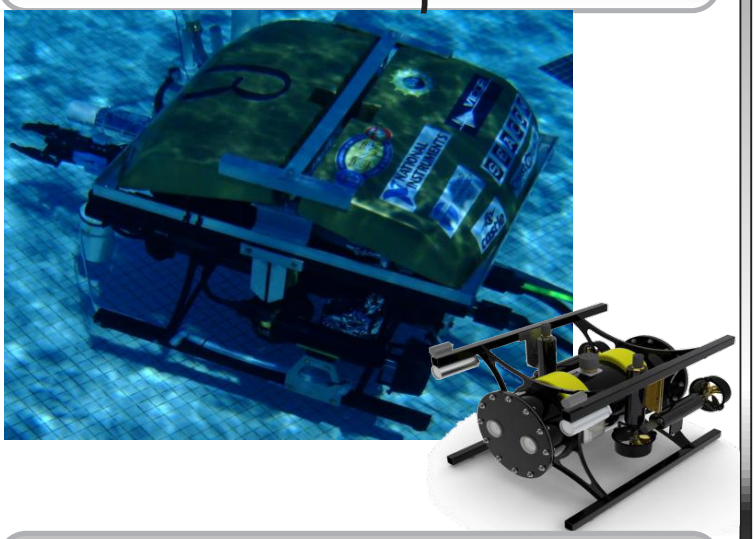
GTSR Remotely Operated Vehicle Fleet

Steven Bradshaw Dr. Fumin Zhang

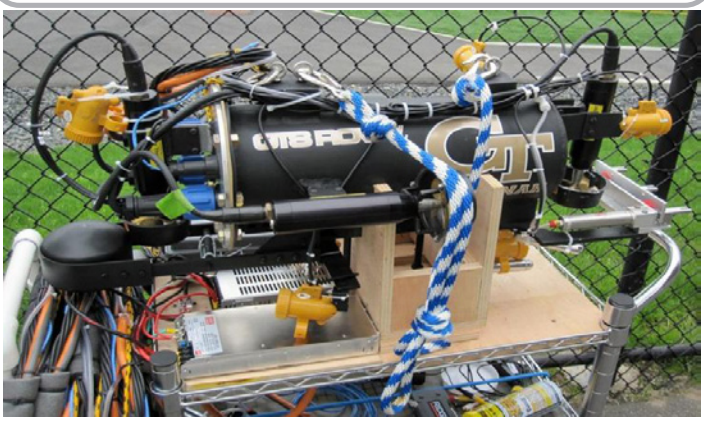
Georgia Tech Robotics & Intelligent Machines @GT

Georgia Tech Savannah Robotics has been researching and developing ROVs since late 2008. Although development was initially pushed by competitions, GTSR's aims are now to develop a versatile underwater platform for research that can also be outfitted for competitions. Undergraduate students gain design experience and practical hands-on experience while developing a platform that can be used for graduate-level research. These student-led projects link research in Cyber-Physical Systems, mobile sensor networks, cooperative exploration and control theory.

ROV-β



ROV-α



Ongoing Research

- Surveying of inland waterways in Louisiana for residual oil
- Field-tracking
- Mapping
- Hardware in loop simulation
- More advanced stability and navigation controllers
- Cooperative exploration
- Vehicle, thruster & battery modeling (for simulation and controller design)
- 3 DOF Manipulation

Competitions

- 2009 M.A.T.E. ROV Competition, Buzzards Bay, Mass.
- 2010 M.A.T.E. ROV Competition, Hilo, HI
- 2011 M.A.T.E. ROV Competition, NASA Facility, Houston, TX

Software Control Systems

GTSR has developed a unified software system which controls all of its vehicles. This enables coherence across vehicle platforms, reduces development time and cost, and increases software versatility. All software is written using LabVIEW 2010.

Team Members

Cooperative Exploration of Level Surfaces of Three Dimensional Scalar Fields

Wencen Wu & Fumin Zhang
School of Electrical and Computer Engineering
Georgia Institute of Technology

Problem Formulation

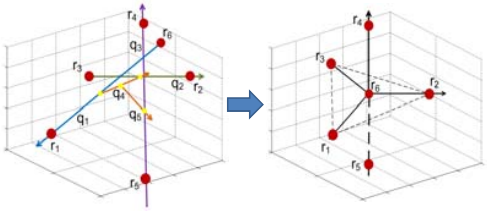
How to explore an unknown 3D scalar field using a group of sensor platforms?

- Cooperatively measure and estimate the field value (i.e. temperature) → Cooperative Kalman filter
- Extract the field information (i.e. principal curvatures & directions) → Curvature estimation using formation
- Track one of the lines of curvature on a desired level surface → Formation shape and motion control

Information Dynamics

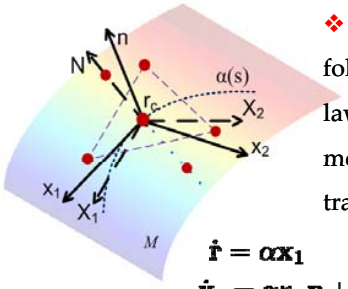
- State: $\mathbf{s}_k = (z(\mathbf{r}_{c,k}), \nabla z_{c,k}^T)^T$
- Measurement equation:
$$p_{i,k} = z(\mathbf{r}_{i,k}) + w(\mathbf{r}_{i,k}) + n_{i,k}$$
- State equation:
$$\mathbf{s}_k = \mathbf{A}_{k-1}^s \mathbf{s}_{k-1} + \mathbf{h}_{k-1} + \epsilon_{k-1}$$
- A cooperative Kalman filter is constructed to combine sensor readings and reduce noise.

Shape Control



- Jacobi vectors are used to describe the formation shape.
- Jacobi transform decouples the dynamics of the formation shape and the dynamics of the formation center motion.

Curve Tracking on a Level Surface

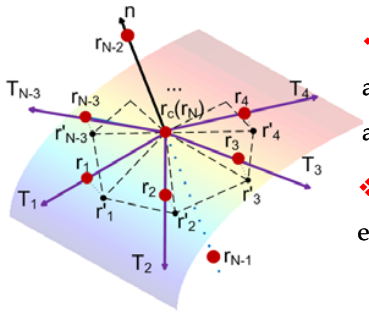


- Differential geometric approaches are followed to develop steering control laws that govern the formation center to move to a desired level surface and track a curve on the surface.

$$\dot{\mathbf{r}} = \alpha \mathbf{x}_1$$
$$\dot{\mathbf{x}}_1 = \alpha\kappa_n \mathbf{n} + \alpha\kappa_g \mathbf{x}_2$$
$$\dot{\mathbf{x}}_2 = -\alpha\kappa_g \mathbf{x}_1 + \alpha\tau_g \mathbf{n}$$
$$\dot{\mathbf{n}} = -\alpha\kappa_n \mathbf{x}_1 - \alpha\tau_g \mathbf{x}_2$$

$$\dot{\mathbf{r}}_c = \mathbf{X}_1$$
$$\dot{\mathbf{X}}_1 = u\mathbf{N} + v\mathbf{X}_2$$
$$\dot{\mathbf{X}}_2 = -v\mathbf{X}_1$$
$$\dot{\mathbf{N}} = -u\mathbf{X}_1$$

Curvature Estimation Using Formation



$$\mathbf{M}_v = \sum_{i=1}^{N-3} \omega_i \kappa_i \mathbf{T}_i \mathbf{T}_i^T$$

- Principal directions are obtained by diagonalizing $\mathbf{M}_v$.

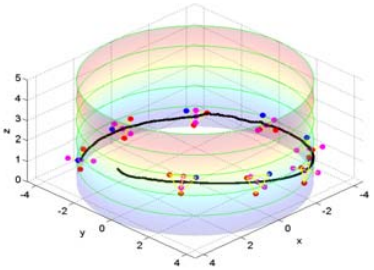
- Taubin's algorithm is modified and applied to estimate principal curvatures and principal directions .
- The algorithm provide nonsingular estimates of principal directions i.i.f:

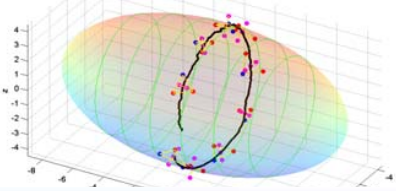
$$\sum_{i=1}^{N-3} \omega_i \kappa_i \sin 2\theta_i \neq 0$$

- $N \geq 6$ must be satisfied to avoid singularity in the estimation.
- $N \neq 7$ if the formation is symmetric.

Simulation Results

- Six sensor platforms are controlled in a desired formation while detecting and tracking one of the lines of curvatures on a desired level surface.





Multiresolution Path Planning Using Beamlets

Oktay Arslan and Panagiotis Tsiotras

School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA

➢ Objective

Develop an alternative representation (graph) of information of the environment which allows to compute paths with curvature constraints

➢ Problem Definition

- Path planning problems in a 2-D environment
- Map size : *n x n* square image
- Black (blocked) and white (free) pixels.
- Problem is to find a path which satisfies curvature constraint between a given pair of source and destination cells.

➢ Multiresolution Path Planning Strategy

Basic Definitions

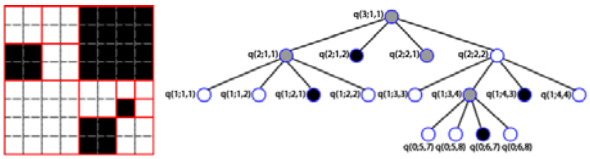
Beamlet : a line segment between any two points

Dyadic square: a set of points which forms a square region.

Multiresolution representation of the environment

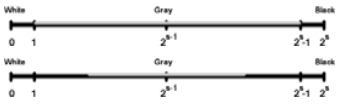
Quadtree decomposition

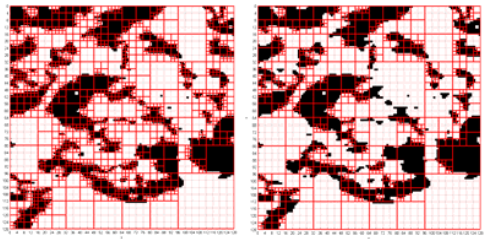
- White, black and gray nodes



- α: tuning parameter for grayness

$$[1+\alpha*(2^{s-1}-1),\ 2^s-1-\alpha*(2^{s-1}-1)]$$





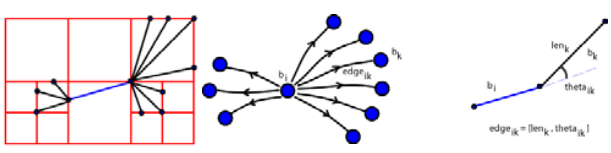
$\alpha = 0.25$ $\alpha = 0.75$

Dyadic partition list

Scale and position information of dyadic square at the leaves of the quadtree

Construction of Beamlet Graph

- Vertex set : all feasible beamlets
- Edge set: two beamlets are connected if
  - they have a common point
  - include both length and angle information



➢ Simulation Results

- Compared with A*, A* PS, Basic Theta*
- Several 2D cluttered environments were created.
  - 2 different gird size (128x128 and 256x256)
  - 4 different percentage of obstacles (%5,10,20,30)
- 100 path planning problems were created randomly

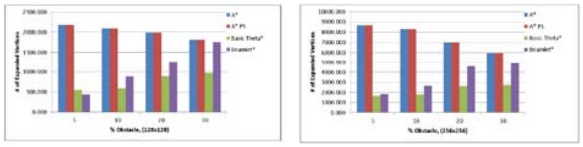
Performance Criteria

l : length of the computed path.
θ_{max} : absolute value of maximum heading angle.
t_e : execution time of the search algorithm.
n_{exp} : number of expanded vertices during search.
n_{hc} : number of heading changes on the path.

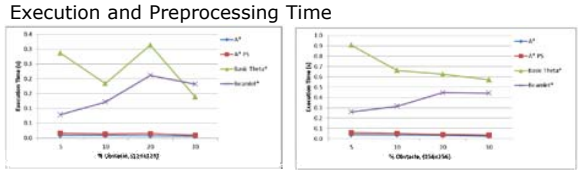
Beamlet* computes smoother paths than other algorithms

	Maps	A*		A* PS		Basic Theta*		Beamlet*	
		l	θ_{max}	l	θ_{max}	l	θ_{max}	l	θ_{max}
128x128	%5 Obstacle	114.351	45.451	110.468	30.071	108.636	15.888	110.574	10.903
	%10 Obstacle	112.873	46.801	109.388	39.372	107.298	22.727	114.188	13.754
	%20 Obstacle	111.753	49.501	108.012	41.409	106.817	31.486	112.627	18.212
	%30 Obstacle	104.305	53.552	101.209	52.710	99.394	44.568	108.859	26.685
256x256	%5 Obstacle	229.153	45.901	223.890	40.205	217.783	17.062	224.676	10.376
	%10 Obstacle	221.885	46.351	217.112	43.631	210.516	22.562	216.837	13.813
	%20 Obstacle	208.623	49.951	203.220	48.556	199.160	36.812	213.703	18.956
	%30 Obstacle	197.481	52.652	192.220	52.397	188.933	42.335	194.818	26.003

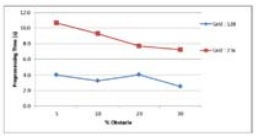
Vertex Expansions



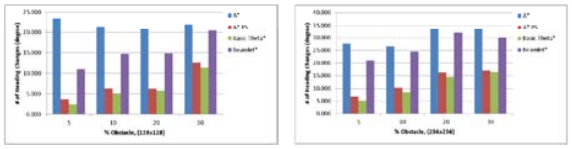
Execution and Preprocessing Time



Preprocessing Time

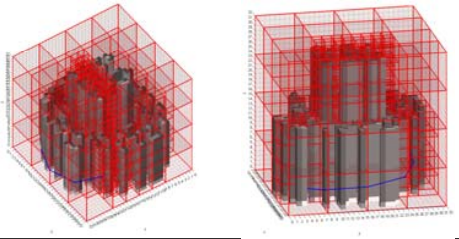


Heading Changes



➢ Future Works

- Extend this approach to path planning problems in 3D



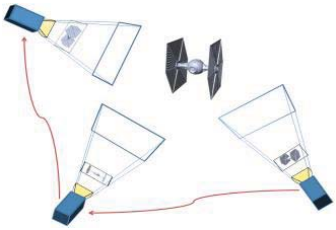
Unit Dual Quaternion-Based Satellite Pose Control for Target 3D Reconstruction

Dae-Min Cho and Panagiotis Tsiotras

School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA

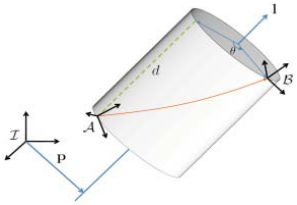
Structure from Motion(SfM)

- Unkown camera pose
- Reconstruct
  - Target(scene) geometry
  - Camera motion



Unit Dual Quaternion

- Simple, easy and compact form to describe rotation and translation at the same time
- Parameterize screw transformation, which can describe any rigid transformation (Chasle's thm)
- Helps design a control law with rotation and translation coupling terms => Better performance
- Better computational accuracy than homogeneous matrix multiplications



$$\hat{\mathbf{q}} = \mathbf{q} + \epsilon\frac{1}{2}\mathbf{q}\otimes\mathbf{t}$$

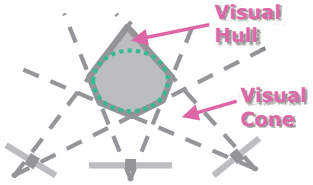
$$\text{where } \epsilon^2 = 0 \text{ and } \epsilon \neq 0$$

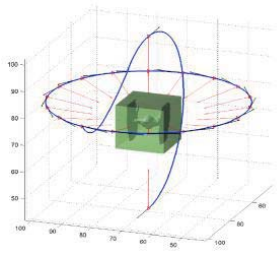
$$\ln\hat{\mathbf{q}} = \ln\mathbf{q} + \epsilon\frac{1}{2}\mathbf{t}$$

- Tracking controller: $\hat{\omega}^s = -2k\ln\hat{\mathbf{e}} + \mathrm{Ad}_{\hat{\mathbf{e}}}\hat{\omega}_d$

Space Carving Algorithm

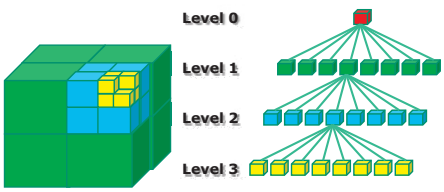
- Silhouette is recursively applied as a mask to remaining voxel array
- Robust to illumination changes
- Visual Hull contains true scene
  - Tightest possible bound

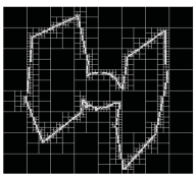




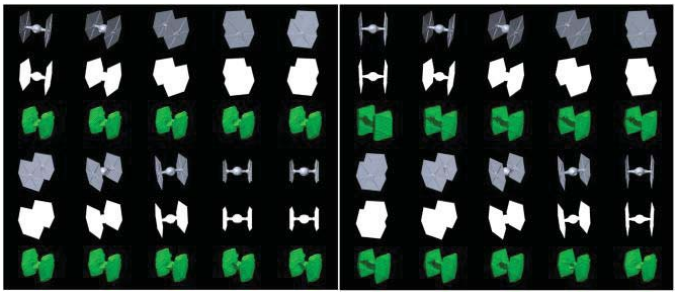
Multiresolution Voxel Representation

- Finest voxels at boundary
- Requires less memory and computation
- Octree represents multi-sized voxel efficiently
- Computationally robust to initial voxel size

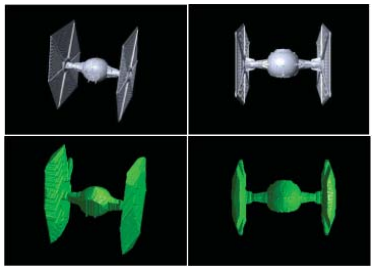




Simulation

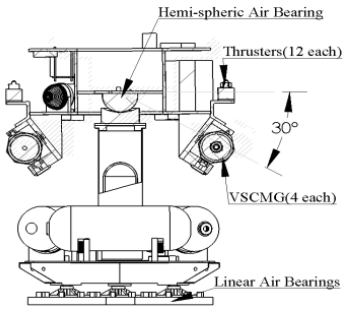


- 20 images are obtained from different camera poses
- Needs to be improved in modeling concave portion



Experimental Testbed

- Designed for Autonomous Rendezvous and Docking experiment in space
- 5DOF(3D rotation + 2D translation)
- Matlab xPC Target based autonomous, real time operation
- CCD camera, 2D Laser Range Finder, and 3D camera
- 12 Thrusters, 4 VSCMGs



Future Works

- Apply unit dual quaternion-based tracking controller to 5DOF Testbed with space carving algorithm
- Space Carving algorithm can be used for inspection and identification of a target

Robotics@GT
& Intelligent Machines

Center for Robotics and Intelligent Machines at Georgia Tech
College of Computing Building • 2nd Floor
801 Atlantic Drive • Atlanta, GA 30332-0280
Phone 404.385.8RIM • Fax 404.385.8546
www.robotics.gatech.edu