Optical Skin For Robots: Tactile Sensing And Whole Body Vision

Summary

Optical imaging will be developed for high resolution tactile sensing. The combination of embedded cameras and soft transparent skin supports measuring contact forces by tracking skin deformation, and tracking nearby objects and surfaces (a ”proximity” sense). The outer layer of the skin is inexpensive and easily replaceable when it wears or is damaged, because there is no wiring or sensors in it. Our approach minimizes wiring, a major source of failure in robots, because the wiring leaving the sensor is already high-bandwidth networking (USB or Ethernet). The wiring, solder joints, and other connections in the sensing system don’t have to repeatedly deform or face cyclic or variable stresses, major sources of sensor failure. The materials that do deform or are stressed (the outer layer of the skin) can be optimized for mechanical robustness. We will evaluate our work by co-developing an ”optical skin” for hands and a synergistic soft hand. Performance on everyday manipulation tasks such as pouring and cutting will be assessed.

Intellectual Merit

A transformative idea is to separate deformation in the skin and its measurement, by developing remote or non-contact tactile sensing. Another transformative idea is to aggressively distribute high resolution imaging over the entire robot body. This reduces occlusion, a major issue in perception for manipulation. Given the low cost of imaging sensors, there is no longer a need to restrict optical sensing to infrared range finders (single pixel depth cameras), line cameras, or low resolution area cameras. A transformation that we hope to lead in robot perception based on skin and other sensors is rich sensing: the fusion of multimodal sensing at a distance, contact-based position, movement, and force sensing, vibration sensing, and volumetric perception of the interior of manipulated soft objects [based on future work on adding ultrasound transducers to the skin].

Broader Impacts

One theory as to why robots are so clumsy is that they have little or no tactile sensing. Endowing robots with superhuman sensing may allow them to be more useful, and cheaper to program or train. We expect to make programming robots to do tasks involving contact much cheaper, and make robots more useful, particularly for care robots supporting everyday life activities. The proposed system is critical for safety of robots that are near or touch humans, and for improved robot performance in many manipulation tasks. Possible applications of this approach include household robots that interact with humans, aware furniture and car interiors, aware clothes, aware tools, and aware floors, walls, and ceilings. Optical skin-based sensing is a practical, manufacturable, and maintainable approach to sensing for touching and helping people, and can be applied beyond robotics to better seats and other furniture surfaces that touch humans, and more intelligent tools and control surfaces that humans touch. A major outreach initiative led by Atkeson is the creation of a physical and virtual Robot Museum. The impact of Atkeson’s work will be increased by a new Disney TV show based on the characters from the Disney movie Big Hero 6, including the inflatable medical robot Baymax inspired by Atkeson’s work on inflatable robots. Atkeson has been successful in graduating a lot of students who are now active participants in the robot revolution.

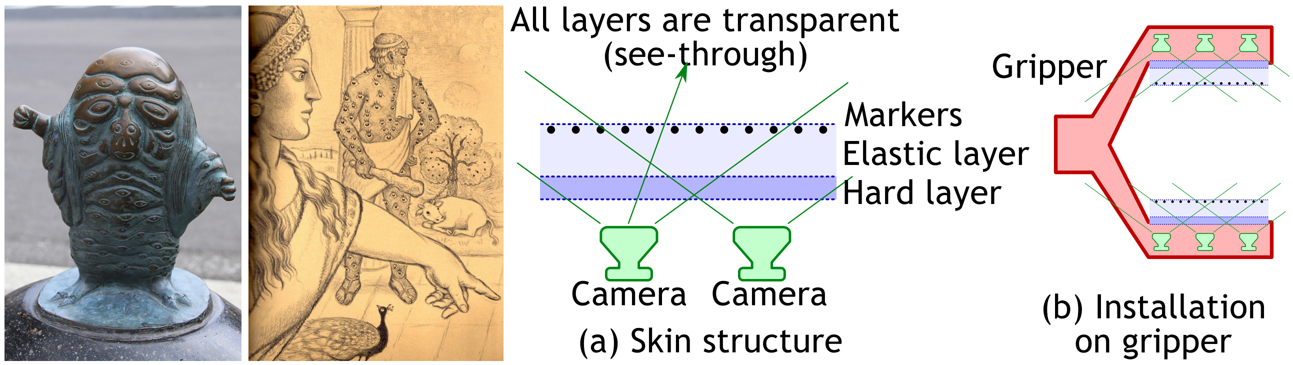

Figure 1: **Left:** *Whole Body Vision* in Japanese (Hyakume) and Greek (Argus Panoptes) mythology. Each figure has 100 eyes spread all over their body. **Right:** Example design of our optical multimodal-sensing skin and an installation example on a robotic gripper.

Optical Skin For Robots:
Tactile Sensing And Whole Body Vision

We propose developing optical sensing and compliant skin for hands for 1) high resolution tactile and force/torque sensing and 2) widespread proximity sensing and tracking of nearby objects and the environment. The *SuperTouch* system will use active multi-spectrum and multi-modal perception to recognize objects and materials, make 3D models of nearby objects and obstacles, detect, localize, and track contact areas and forces before (prediction) and during grasping and other forms of contact, measure and control forces, monitor object and hand deformation, tearing, and other damage to manipulated objects, and predict, detect, and measure slip.

Motivation

Robots currently avoid contact, rather than use it to make tasks easier. In the DARPA Robotics Challenge (DRC) Finals, no robots used railings, walls, door frames, or obstacles to guide or support the robot body [10]. More research in control algorithms is needed, but a major problem for most DRC teams was the lack of tactile and proximity sensing over their entire robot body. There were many sensors in the head, but often the robots could not see contacts. We added wrist and knee cameras to our Atlas robot to watch the hands and feet. Even this sensing was not enough, and the bulky cameras got in the way.

Robots need skin with appropriate mechanical properties, rather than just using hard shells. We need robot skin that is compliant rather than stiff, and has appropriate viscosity, friction, hysteresis, and other mechanical properties. Skin often needs to be applied to doubly curved surfaces and with various thicknesses and stiffnesses. This makes it difficult to use flexible circuit boards and other stiff flexible sheets, because they don't stretch to conform to a doubly curved surface. We are also interested in soft sensors for inflatable robots (Figure 5).

A major motivation for this work is our experience with currently available tactile sensors and robot hands, and considering the life history of robot skin. Skin is repeatedly subject to deformation, wear, and other damage. Materials harden, soften, deform (sag), or become brittle with age or exposure to sunlight. Fluids (gas or liquid) diffuse, leak, separate, and evaporate. This is increased if the fluid containment is repeatedly deforming. At least the outer layer of skin needs to be easy and inexpensive to refill, repair, or replace. Humans replace their skin outer layer approximately monthly. In robots, materials that do deform (the outer layer of the skin), should be able to be optimized for repeated deformation as well as desired mechanical properties, rather than for sensing.

It is important to minimize the maximum stress and resulting deformation and the stress and deformation

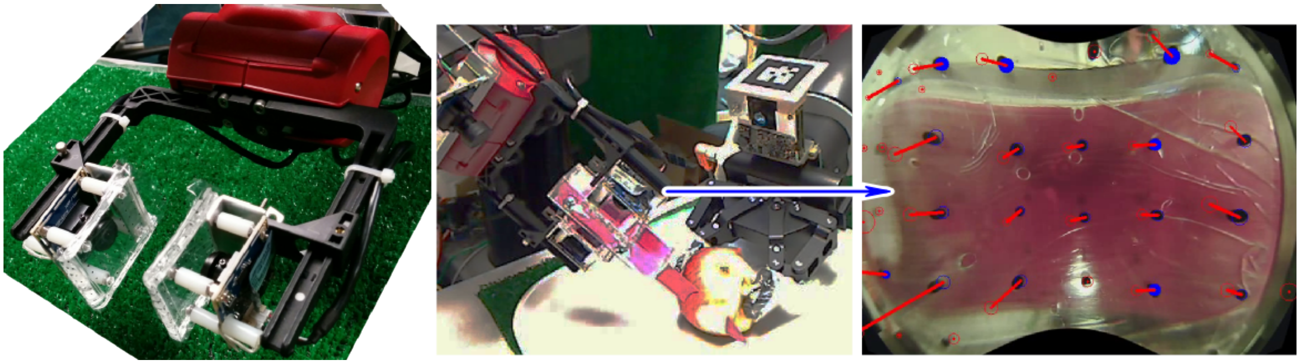

Figure 2: FingerVision: Prototype of the optical skin on the fingers of the Baxter robot. Left: the sensing skin and the fingers. Middle: cutting an apple. Right: a camera view during cutting where the marker movements are rendered.

cycling of wires, electrical components, flexible circuit boards, and solder joints. Even wiring designed to stretch and deform (for example coiled wiring embedded in elastomers) fatigues and fails relatively quickly. This suggests favoring sensing designs that do not put electronic components or wires in the part of the skin that deforms.

In general we want to minimize the number of components and the amount of wiring in terms of increasing the system probability of failure as well as costs associated with these elements. This suggests favoring designs that have few components. One could argue that the imaging chips we propose to use have millions of components, but overall failure rates have been brought down to acceptable levels, so we can treat a camera as a single component.

Implementing compliance is difficult for tactile sensors with rigid components such as rigid planar printed circuit boards and micro-electro-mechanical systems with planar substrates, so we want to avoid these in the deformable region. Flexible PCBs can be bent only in one direction at a time unless cut, and are not stretchable. Their mechanical robustness to repeated deformation is limited. Flexible PCBs are more robust if adhered to a shaped rigid substrate.

Most tactile sensors, and especially tactile sensor arrays, only measure normal force. Tangential forces are especially important for predicting and avoiding slip.

Our goal is to complement existing tactile sensor approaches as much as possible. Our proposed approach, which is limited to frequencies defined by frame rates, complements and combines with high frequency vibration and texture sensing provided by embedded inertial measurement units (IMUs: accelerometers, gyros, magnetometers, and temperature sensors), microphones, and other vibration sensors. We will also explore the use of hairs and whiskers to extend tactile sensing's range [42]. We note we do not have to instrument the hairs and whiskers, as they can be optically tracked at a distance. This makes hairs and whiskers inexpensive to use. We note that optical mice use sensors with frame rates in the tens of thousands, but we have not found a way to transfer these frames out of the chips (such as the ADNS-9800) at more than 100fps, which is about what we get from our other cameras. Total image shifts for each frame can be transferred at much higher rates from these specialized cameras that compute optic flow.

The robot hands we have worked with (Figure 7) have had fat fingers with hard surfaces and rigid palms. We will explore soft hands with internal skeletons and flexible palms that are designed with sensing in mind, rather than treating sensing as an afterthought in hand design, and assuming thin sensor pads can just be added to a bulky shell.

The most recent workshop on tactile sensing (yesterday), at the 2016 IEEE-RAS International Conference

Figure 3: Kinematic prototype of a finger skeleton with cameras.

on Humanoid Robotics, identified the following issues in tactile sensing: "1) Available tactile sensors are still far from matching the capabilities of the sense of touch in humans; 2) Tactile data processing algorithms and data-driven learning technology are still immature; 3) A general and robust robot manipulation control framework is still missing." [62]. This proposal addresses issue 1, and enables work on issue 2. Other work in our lab supported from other sources addresses issue 3, so we will be able to address these issues in an integrated way.

Scholarpedia's article on tactile sensing says: "Optical sensors are attractive due to their potential for high-spatial resolution, robustness to electrical interference, light weight, and their potential to resolve the wiring complexity problem." and "Tactile sensing systems are still relatively undeveloped compared to the sophisticated technology accomplished in vision. The relatively slow development attained thus far is possibly related to the inherent complexity of the sense of touch. Another limiting factor is that, by their very nature, tactile sensors require direct contact to be made surfaces and objects, and are therefore subject to wear and risk of damage than some other sensor types." [58].

The German lab DLR states: "Requirements [for tactile sensing] include robustness regarding wear and tear, easy low-cost mass production and a space-saving, fast read-out mechanism. The skin should be soft and exhibit mechanical properties comparable to that of human skin." [19]

We do note that the field of tactile sensing has been promising, but not delivering, for decades. Rob Howe's web page states: "Thirty years of benchtop research have failed to deliver tactile sensing suites that are robust and inexpensive enough to provide grasping information at a practical cost/benefit ratio." [28]. The 2016 Springer Handbook of Robotics states: "Tactile sensing always seems to be a few years away from widespread utility." [17].

Finally, here is an interesting quote: "One reason for the slow development of tactile sensing technology as compared to vision is that there is no tactile analog to the ... CCD or ... CMOS optical array." [17]. Perhaps it makes sense to just appropriate CCD and CMOS optical arrays from the vision folks and adopt them for tactile sensing as well.

Related Work

Figure 4 shows some examples of tactile sensor design paradigms. Recent reviews of tactile sensing are provided by [17] and [18]. Recent European projects include TACMAN: Tactile Manipulation and Roboskin. The literature is huge and many prototype sensors have been built. It is not possible given the page limits to survey non-optical tactile sensing, so we restrict our related work survey to optically-based tactile sensing. Few tactile sensors are actually widely used, and they are restricted to very simple designs. Capacitive sensing and other contact tracking methods for track pads and touch screens are a success, but not directly applicable to robot skin. The use of simple optical sensors (range finders, for example), switch contact detection, load cells, force sensitive resistors (FSRs) and capacitive and resistive array sensors probably dwarf all other types of industrial (and robot) tactile sensors.

Here are the goals that lead us to focus on optically-based tactile sensing:

- We want tangential forces as well as normal forces, because slip is important. This criterion eliminates

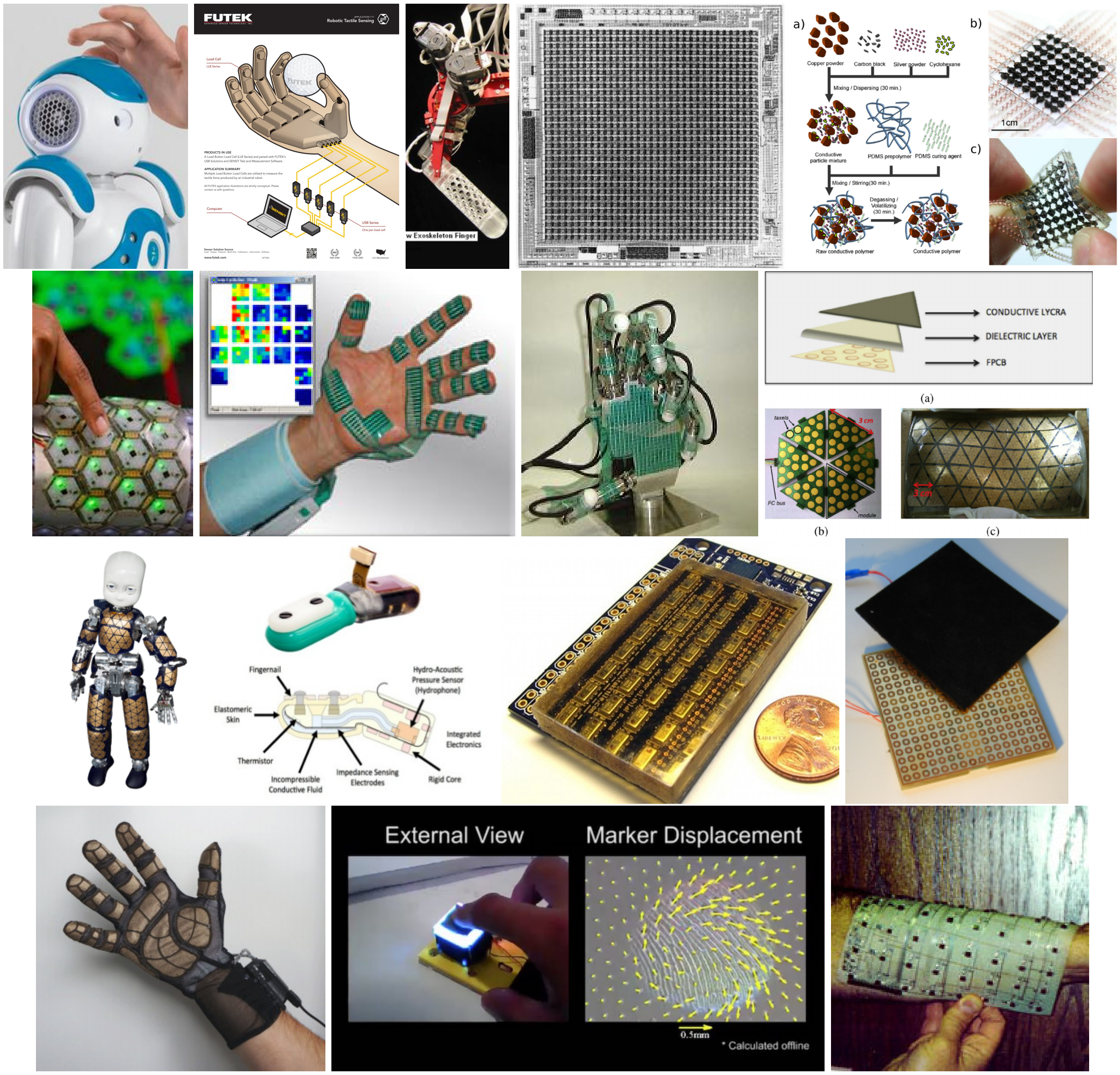

Figure 4: Some examples of current tactile sensor design paradigms. **L-R:** [NAO] Plates or pads instrumented with switches or capacitive sensors such as those used in touch screens. [Futek] Single or multi-axis load cells at fingertips or at every possible contact point. [StanfordHaptics] Resistive strain gages, fiber optic strain sensors, or other strain sensors embedded in structural elements such as finger links. [Toyota] Circuit and MEMS sensor arrays. [NTU] Discrete elements printed or embedded in an elastomer which measure local properties, or relative positions, connected by printed or embedded wires. [TUM] Rigid printed circuit board (PCB) tiles. [Tekskan] [Gifu] Circuit elements on flexible PCBs, which can measure resistance, capacitance, or inductance changes. Currently force sensitive resistors (FSR) are the most popular circuit elements. [iCub] Arrays of flexible PCBs. [BioTac] Bags of fluid (gas or liquid). In addition to pressure, electrical or magnetic field measurements can measure deformation. [Takktile] Elastomer with underlying discrete pressure sensors mounted on a rigid PCB. [Bielefeld] Soft sensor arrays. [Bielefeld] Fabric based sensing. [GelSight] Optical measurement of elastomer deformation (similar to proposed approach). [Lumelsky] Distributed infrared range sensors. Paradigms not illustrated include MEMS approaches, approaches which embed materials or wiring in an elastomer and measure bulk resistance, capacitance, dielectric properties, or inductance, conductive liquid, thin metal film, piezoresistive, piezoelectric, and microfluidics plus capacitance measurement.

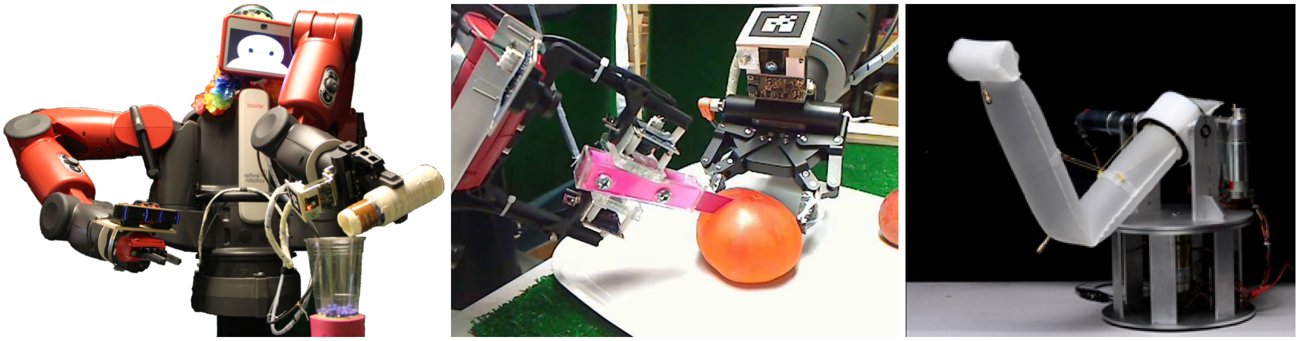

Figure 5: Our Baxter robot, Baxter cutting a tomato, and an inflatable robot.

most potential tactile sensing technologies.

- We want high resolution. “In comparison to animals, with hundreds or thousands of mechanorecepters per square centimeter of skin, even the most sophisticated robots are impoverished.” [17]. This criterion also eliminates most potental tactile sensing technologies.

- We want immunity to electromagnetic interference. This favors optical approaches over capacitive and magnetic approaches.

- We want to minimize the cost and embedded components in the portion of the skin that undergoes deformation or stress due to contact. This goal is met by optical systems and electric and magnetic field systems, as well as some exotic approaches that effectively use sonar or ultrasound to do tactile sensing.

- We would like a proximity sense. Again, this goal is met by optical systems and electric and magnetic field systems, Interestingly, optical systems provide a proximity sense when there is no occluding contact. Sonar and ultrasound approaches provide a proximity sense where there is a contact, and the lack of contact is occluding.

- We want to minimize wiring and component count. Since small cameras and their high speed networking are highly developed, it is easy for camera-based systems to meet this criteria.

The idea of *Whole-Body Vision* is thousands of years old (Figure 1). We want to go beyond palm, wrist, crotch, and knee cameras to cover the entire robot with cameras. Palm and wrist cameras are in common use and are commercially available (examples include [67, 68]). A typical home-brew wrist camera is shown in Figure 5. Using a large number of cameras has been pursued in the development of VR cameras, where tens of cameras are mounted on the same rigid body. We want to put more cameras on multiple moving parts of a robot.

The idea of using imaging sensors for tactile sensing is decades old. An initial attempt was measuring the frustration of total internal reflection within a waveguide on a sensor surface caused by contact [13, 57, 101, 34]. The research trend has shifted to measuring displacement of markers placed on the sensor surface with computer vision [39, 83, 16, 29, 35, 36, 6, 49]. The resolution of the measurement is determined by the camera and the marker density. Recently high resolution sensors have been proposed (e.g. [51, 49]). Another idea is markerless localization of a registered object (e.g. [51]). The localization accuracy is ideally equal to the camera resolution. Much of the above work uses a transparent elastic material between the sensor surface and the base. The dynamic range of the force measurement can be controlled by changing the hardness of the elastic material (softer is more sensitive; cf. [65]). Many have proposed the combination of structured light, markers, and multiple imagers to estimate deformation in transparent skin [74, 64, 61, 94, 30, 40, 41, 82, 93, 72, 27, 71, 1, 33, 45, 14, 37, 50, 100,

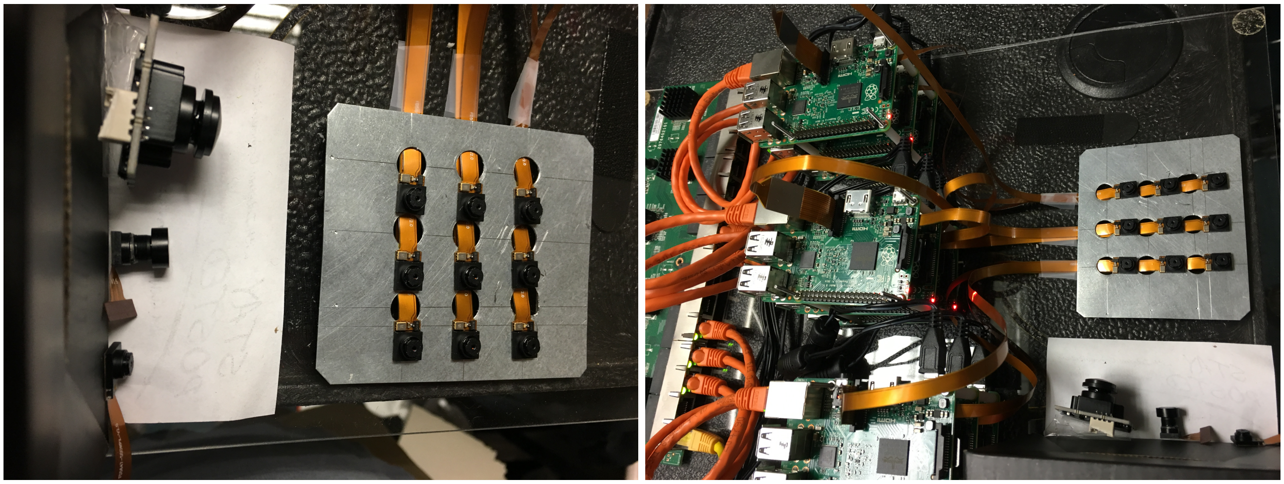

Figure 6: A 3x3 camera array using cell phone cameras. The cell phone image sensors used do not perform image compression, so substantial additional computer hardware was added for image compression and networking. This is unnecessary for the larger camera used in *FingerVision I.* shown on the left and bottom of the two images for size comparison, since it has onboard image compression.

65]. Saga et. al. added hairs to their optically sensed skin [69]. Our research is close to these approaches. An important difference is the total transparency of our skin including the outer surface, which gives us vision of external surfaces as well. Previous work used an opaque top layer on the skin to block external light as it would affect marker tracking. We will solve the marker tracking problem under natural external scenes using computer vision in order to make use of totally transparent skin.

Embedding optical range finders (single pixel depth cameras) in the skin of a robot to provide whole body distance sensing was proposed more than 40 years ago and continues to be developed [38, 84]. Another sensor with fully transparent skin is proposed in [66] where infrared range finding sensors are used instead of a camera. The idea is to measure distances between the sensors and an object, and estimate the deformation of the transparent skin from the distance. Vertical contact forces are estimated from the deformation. If there is no contact with an object, this sensor simply gives the distance to a close object. We believe imaging can be used instead of just range finding for great benefit and not much added cost. Although this sensor and ours have different sensing modalities and ranges, we can share ideas; e.g. we could embed distance sensors around the cameras.

Space limitations prevent us from covering all the relevant background in visual and tactile perception; reasoning about contact; trajectory, force, and contact control; reinforcement learning; discriminant or predicate based policies; matching learned sensory templates; deep learning, and robot learning in general, Rather than give a series of superficial reviews, we focus here on signature learning, detection, and verification. We note the increasing popularity of "kitchen sink" or "black box" machine learning approaches to contact and object tactile recognition and tracking [47]. One popular machine learning approach is neural networks (deep learning) [63].

The data mining community has developed several varieties of algorithms for mining discriminative key signatures from various types of multimodal data. Shapelet discovery [99] finds discriminative subsequences of time-series data, but has been generalized to other forms of data like classifying plant species based on leaf outline shape. It can also learn from multimodal data by learning decision trees, combining subsequences from multiple modes of data into a single classifier. Chiu et al. [15] detect repeated time series "motifs" with a probabilistic anytime algorithm that utilizes random projections. SAX [52] is an alternate approach to classifying time series data by using a discretized symbolic representation of the data that is conducive to dimensionality reduction. More recently, spatio-temporal hierarchical matching pursuit [56] has been used to to learn time-varying features from tactile data to classify grasp stability and to perform object recognition from

Figure 7: Hands Atkeson has worked with. Note the prevalence of hard surfaces and rigid palms. **L-R:** Utah/MIT, Barrett, IRobot, Robotiq, and Sandia hands.

touch alone. McManus et al. [59] introduce an algorithm to detect *scene signatures* in image data that help to recognize specific physical locations under many different weather and lighting conditions.

Preliminary Work

To develop our ideas, we created FingerVision: a tactile and proximity sensor for parallel jaw robot grippers [96]. This video of the system is a good introduction to our work [95]. Figures 2 and 5 show a prototype of FingerVision I and its installation on our Baxter robot fingers. We demonstrated the usefulness of the optical skin in cutting food. Figure 6 shows a prototype of a 3x3 camera array using cell phone cameras, where we tested networking and processing nine compressed video streams.

Proposed Work

Key issues in the proposed work include: 1) Developing a skin prototype that can be used to develop perception algorithms and measure computation and networking usage. This prototype will put approximately 100 cameras on the shell of our Baxter robot. 2) Developing a hand with skin to evaluate the proposed system on tasks, and to assess the mechanical robustness and lifetime issues of the proposed system. 3) Exploring how to use multi-spectral, multi-modal and active sensing to improve tactile and proximity perception, especially focused on human detection and localization at the human skin level, activity recognition, and contact management. 4) Developing tactile and proximity perception algorithms for this system. 5) Developing a better low profile camera and lens system suitable for the skin.

Proposed Work: Whole-Body Vision Prototype I

We propose building a prototype of the *SuperTouch* system with a hundred optical sensors such as visible light (SONY IMX219 8-megapixel sensor, for example) and near and far-infrared (thermal) image sensors (Figure 11), camera modules and board cameras, GigE CAMERAS, and RGBD imagers. This prototype would have a gigabit Ethernet network and embedded processing (perhaps using NVIDIA Tegra processors) to compute optic flow on single camera streams and stereo matching on image streams from pairs of cameras.

Based on our previous work with *FingerVision* and the network shown in Figure 6, we will develop a whole-body vision system for our Baxter robot. This system will be used to develop perception algorithms and measure computation and networking usage.

We estimate it will involve approximately a hundred cameras on the shell of our Baxter robot. This is about an 11x replication of our 9 camera test setup, so we are confident we can build the hardware. In fact, it will make an excellent undergraduate project for our REU student. We estimate the cost at less than $100 per video

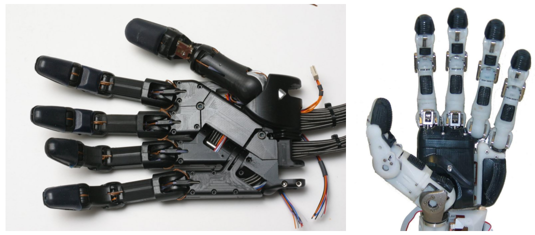

Figure 8: Hinged palms. **L-R:** Shadow and Elu2 hands.

stream (including camera, cables, and a share of aggregating computers and post-aggregation network cables and switches), so the cost is comparable to hiring a summer undergraduate student. Initially, we will route the video streams to an existing local computer cluster for processing.

Proposed Work: Co-development of Hand and Sensing

Our goal in this section is developing a hand with skin to evaluate the proposed system on tasks, and to assess the mechanical robustness and lifetime issues of the proposed system. Figure 3 shows a kinematic prototype of a skeleton for a finger with cameras embedded in the skeleton. We will also explore embedding cameras in soft skin material. The finger is not yet actuated. Tendons will be used to flex the finger joints, and will be routed to the side of the cameras. Flexure joints will extend the finger. Initially Dynamax servos will be used to actuate the tendons.

We will take sensing into account when designing hands (sensing seems to be an afterthought in most hand designs). Figure 7 shows hands Atkeson has worked with. These hands are characterized by hard surfaces, rigid palms, inadequate sensing, and frequent mechanical failure.

We will explore skeletal hands, where an internal skeleton is covered with soft material. A key issue we will explore is non-rigid palms. Some hands have been designed with hinged but otherwise rigid palms (Figure 8). We will consider hinged palms designs, as well as continuum designs where the palm can bend in many ways and does not have a small set of fixed joints. We will also consider hand designs that extend the fingers into the palm, and cover the “metatarsal bones” with fleshy material (Figure 9). Figure 10 shows a kinematic prototype of a skeleton for a hand. The key point is that the palm can conform to grasped objects due to the use of metatarsal bone structure. We will use similar actuation for the full hand as what we plan to do for the finger (tendons to flex, flexures to extend, and remote servos to pull on the tendons).

Proposed Work: Multi-spectral, Multi-modal, and Active Sensing

Previous skin and tactile sensing projects typically focused on one or only a few types of sensors. We will deploy many types of sensors. in addition to using visible light optical sensors to measure skin deformation. Electrical properties of the skin including resistance, capacitance, and inductance can be measured. Printed antennas and inductive coils similar to what are used in wireless RFID anti-shoplifting devices may also be usefully placed on the skin surface or embedded in the skin to measure static and dynamic electric fields, as well as magnetometers or Hall effect sensors to measure magnetic fields. Radar chips are being developed for monitoring respiration and hand gestures at a distance [25]. Pressure sensors can be used to measure skin forces. Given the low cost and small size of far infrared (thermal) imaging sensors, there is no longer a need to restrict sensing to just visible light, or just near infrared. For robots that work with people or processes involving changing temperature (e.g. cooking) imaging in the infrared spectrum is useful (Figure 11), as well as skin temperature sensors. Small time of flight depth cameras are now available (Figure 12). Ultrasound transducers

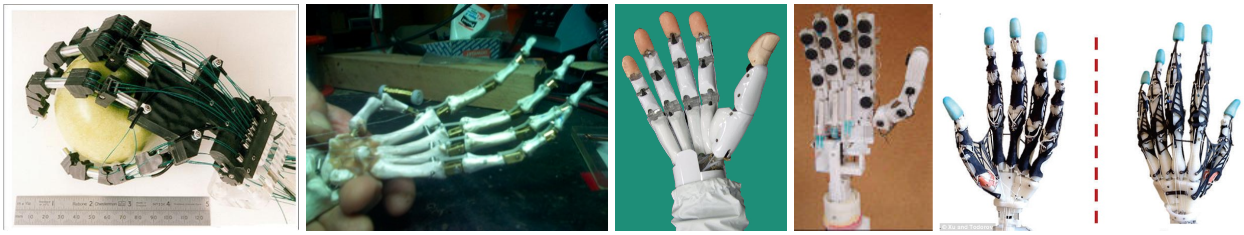

Figure 9: Hands with metatarsal "bones". **L-R:** Sheffield, Lara, Miyazaki Lab, OCU, and Todorov hands.

.

can be built into robot skin to image objects and human tissue that are in contact to avoid damage, injury, and pain. Accelerometers, gyros, IMUs, piezoelectric sensors, and microphones are useful to detect vibrations, localize contacts, and recognize texture and material properties [75, 32]. Accelerometers are also useful to measure orientation relative to vertical (given by the direction of the gravity vector). High speed imaging used in optical mice (essentially using very high frame rate cameras with low angle of incidence illumination (Avago ADNS9800, for example)) can detect horizontal skin, object, and environment movement. Hairs or whiskers glued to piezoelectric sensors or optically tracked provide mechanical sensing at a (short) distance. It may also be possible to embed mechanical elements in the skin that click or rasp when deformed, and use microphones to track skin deformation. We will explore deliberately creating air and liquid (sweat) flows (both inwards and outwards) for better sensing (measuring variables such as pressure, conductivity, and temperature) and controlling adhesion. We will explore humidifying the air for better airflow sensing, contact management, adhesion control, and ultrasound sensing.

We will use joint angle measurements, augmented by link angular velocities, linear accelerations measured by MEMS gyros and accelerometers, and *Whole Body Vision* to estimate where the robot parts are. This state estimation is improved by and improves skin-based state estimation of robot, object, and environment configuration.

Another goal is to explore multimodal optical perception as well as the extensive use of active structured light. We will need to at least add LED-based illumination for when outside light is inadequate or occluded. Internal illumination can be used to enhance texture perception. Structured light sources can be used to improve the localization of markers embedded in the skin and nearby surfaces. In some sense infrared range finders and Time of Flight (TOF) imaging depth sensors use structured light, as does the first version of the Kinect. We can use similar structured light projection to improve surface reconstruction and tracking measurement. It is possible to even use small DLP projectors embedded in the skin to vary the structured light pattern according to viewing conditions [46].

Proposed Work: Algorithm Development

The key features to be demonstrated by this prototype are: 1) surface reconstruction of nearby surfaces, 2) contact prediction, 3) contact area tracking, 4) normal and tangential force estimation over each contact area, and 5) attention-based control of processing to allocate processing resources efficiently.

In the proposed research we will focus on finding out how far we can go with publicly available vision software such as OpenCV and other software. For example, SLAM software for multiple cameras is available from local CMU colleagues as well as other groups such as the TUM Computer Vision Group. We will integrate existing software modules rather than write new low level vision algorithms. Future research will focus on developing new algorithms for optically-based tactile and proximity sensing.

We will explore a range of perceptual approaches, including object tracking based on contact types, forces,

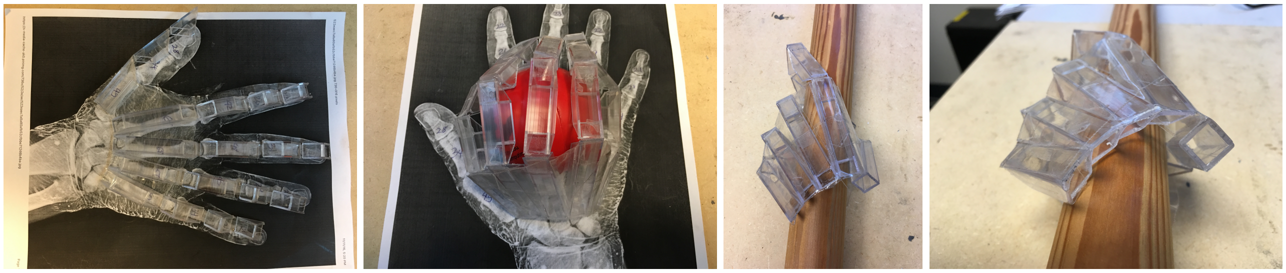

Figure 10: Kinematic prototype of hand skeleton. The palm can conform to grasped objects due to the use of a metatarsal bone structure.

and distances, feature based object recognition based on features such as texture, stiffness, damping, and plasticity, feature based event recognition based on spatial and temporal multimodal features such as the frequency content of vibration sensors, and multimodal signature based event recognition.

As an example of new opportunities that open up with a rich set of sensors, we will discuss event recognition and physical interaction tracking based on "signatures." We will use our rich set of sensors to identify signatures of key events, such as insertion of a peg or screw in a hole. Our embedded cameras might be able to see the peg and/or the hole. Scraping the peg across the surface may lead to an impact and resultant shock wave that is picked up by vibration sensors when the peg is dragged across the edge of the hole. In previous research in our group in collaboration with a postdoc, Scott Niekum, we have conducted a preliminary exploration of using data mining techniques [99] to discover *key signatures* of success and failure from multimodal data from sources such as microphones, wrist accelerometers, and RGB-D cameras. For example, the co-occurrence of a clicking noise with an accelerometer spike may reliably indicate success, while either feature on its own might be insufficient (an accelerometer spike could be caused from sliding off the table, or a click from missing the hole and hitting the table). With these features alone, we have achieved a classification success rate of around 85%, but we suspect that this will greatly improve with more sensory modalities such as pressure and shear force. This same logic applies to more dynamic tasks as well, such as pancake flipping or other manipulation tasks during cooking, which rely on making control decisions based on friction, stickiness, and slippage.

How can a robot know when success has been achieved or if various types of failure have occurred? An ongoing area of our research explores the use of data mining and deep learning techniques to learn how success and failure manifest themselves across multiple perceptual modes such as visual, audio, and proprioceptive data. The classification of multimodal time-series observations is complicated by several factors. First, any given example of success or failure may appear substantially different than other similar outcomes due to the particular configuration of the world or the actions of the robot. Second, a large fraction of the robot's observations during the execution of a task may contain no information about success or failure. Third, success or failure in a task may not be able to be determined from execution data alone, requiring additional interaction with the system.

For these reasons, part of our work searches for *key signatures* of success and failure—discriminative features mined from subsequences of data. This approach allows the robot to use relatively large amounts of data in a semi-supervised fashion to characterize important subsignals, while ignoring irrelevant sections of data. Rather than look only at execution data, we are also exploring teaching strategies through demonstrations that allow the robot to actively gain data with greater discriminative value. Finally, we are also investigating unsupervised learning techniques to identify different types of failures so that appropriate recovery strategies can be learned for each failure mode.

Our *SuperTouch* system will generate a flood of information, and we will have to distribute computational resources throughout the robot to collect and distill the flood, so that only parametric information and surprises are forwarded to "higher" levels of processing.

One way to reduce the information flood is to have local processing at the camera compute parameters or

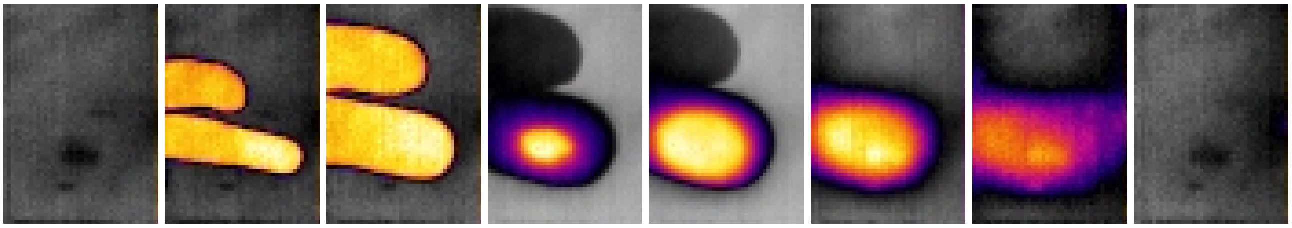

Figure 11: A time series of far infrared (thermal) images from a camera looking through a skin at a finger touching the skin. The skin is transparent in the far infrared spectrum, as well as for visible light. Due to the large dynamic range of the sensor, each picture is scaled so the hottest value is yellow. Note that it is easy to tell which finger actually touches the skin, that the skin is heated up by contact (very quickly), and there is an afterimage as the skin cools off. The camera used was a Lepton LWIR module [24].

sufficient statistics of what is seen, and only forward parametric information such as optic flow parameters or blob tracking results. With local processing we could do local computation of optic flow and stereo disparity of adjacent cameras. Similar computations are done for camera stabilization and image compression, which are available now on many imaging chips. We could locally compute changes to previously computed local surfaces. We could build surface models of what a local set of cameras are seeing.

Another way to reduce the information flood is to not process each image stream fully all the time. We expect attentional processes to determine what processing is applied to which image stream. Attentional processes can select which video streams and which frame rates, frames, and/or regions of interest (ROIs) from each video stream should be processed. We expect one type of attentional process to be applied using local processing at the camera, to detect "salient" events, such as large spatial discontinuities in illumination or motion. Parametric predictions will be supplied to local processing at the camera so that "surprises" can be detected.

We believe interaction and internal sensing on a complex active platform such as a robot must be model-based, and that the models will have to be learned (data-driven). The model determines the expected sensory readings, and what is perceived is derived from the difference between the expected and actual sensor readings. A research challenge is to find ways to use inaccurate sensing and imperfect models to build improved models and more accurate sensing. Another research question is how to most effectively utilize our sensing systems, such as deciding what to look at with our *SuperTouch* system. We will pose this as an optimization problem, taking into account the energy used to drive active sensing systems, the cost of resources used to process the raw sensing data, and the anticipated utility of the information gain.

Early versions of this system would simply move all the video streams to a cluster of computers with multiple GPUs. Later systems would put more computation nearer to the cameras, such as the low power (10W) NVIDIA Tegra CPU+GPU systems. Our larger cameras have on-board MJPEG and sometimes H264 compression. Assuming a high quality video stream is 5Mbps, a single Gigabit Ethernet line can handle on the order of 100 cameras, and a 10G Ethernet line can handle 1000 cameras. Another concept we can use is compressing N related (similar) video streams together (ensemble compression) for further compression.

Proposed Work: A Better Low Profile Camera

A key road block to developing small cameras is developing low profile lenses. Commercially available lenses are often too large or have a too narrow field of view. We propose to develop appropriate lenses for whole body vision. Unfortunately, it is difficult to directly 3D print optically clear lenses on most 3D printers, because each added layer or voxel creates a material interface with imperfections and potentially trapped air bubbles. One approach is to mold lenses and use CNC machining or 3D printing to create the molds [76]. Lens designs can also be commercially 3D printed using a special process that produces optically clear material [55]. Figure 12

Figure 12: **Left:** A small time of flight camera. **Middle:** Front and back view of a lens for an optical mouse, and faceted dome lenses (transparent to infrared). **Right:** Shisha embroidery, in which mirrors are sewn into fabric. We propose sewing cameras into fabric. Low profile lenses are needed for this approach to whole body vision.

shows some lenses we will use as starting points for our designs: the lenses used in optical mice, and faceted dome lenses, in which each facet is a flat Fresnel lens. Faceted dome lenses are in some sense intermediate between a single lens system and separate lenses for each pixel found in a compound eye. We will explore compound eye designs as well, using different lenses for different portions of a single imaging chip. This turns one high resolution imaging camera into many lower resolution cameras pointed in different directions.

Another approach we will explore is lens-less cameras. One example is FlatCam (co-developed by our colleagues at CMU) which uses a mask rather than a lens [5]. We will also explore combining compound eye designs, lenses, and masks. A mask can be thought of as many pinholes. Faceted dome cameras can be thought of as 3D surfaces with many holes, with a lens in each hole.

Research Plan:

This research involves one postdoctoral fellow, in addition to the PI. The PI will supervise, guide, and provide assistance as needed, as well as focus on theoretical and algorithm development and building device prototypes.

Year 1: The major emphasis of year 1 will be building a sensing skin for our Baxter robot, and perception algorithm development. We will do preliminary work on designing the proposed hand and hand sensing system. We will do exploratory development of new lens systems.

Year 2: Year 2 will focus on co-designing and building the hand and hand sensor system. Our design will be informed by the results of the Baxter sensing skin developed in Year 1.

Year 3: Evaluation and refinement of our approaches and low profile camera development will be the major themes of the third year. We will evaluate the approaches both from an experimental and a theoretical point of view. We will evaluate and improve our hardware implementations, and identify and resolve new issues that result. In the third year the goal is to demonstrate our sensing Baxter with our hands performing and learning a wide range of tasks.

What Might Go Wrong?

There are several possible reasons this approach might turn out not to be useful. 1) The first challenge is space and geometry. Can we fit cameras in useful places? Can we give them enough of a view of the skin, and through the skin, to be useful? Wire routing can go behind the cameras, but tendon routing often cannot, due to moment arm considerations. If camera coverage is incomplete, is it enough to be useful?

2) The second challenge is processing. Can we get enough processing power to do something useful? Given that optic flow and stereo disparity can be currently calculated using camera stabilization and data compression ASICs and FPGAs, we are confident the answer to this question is yes.

3) The next challenge is power and cooling, which do not have to be solved immediately, but will have to be solved to develop an autonomous system.

4) The last challenge is unique to optical systems. The robot skin will get dirty and dusty, and it may be necessary for the robot to implement self-cleaning behaviors, or be cleaned regularly by humans.

Given these risks, the proposed research is still worth doing because we need to quantitatively find out whether the above questions can be resolved favorably or not. We won't have accurate estimates of space, processing, power, and cooling needs until we build a prototype and try it out on example tasks. We don't believe accurate estimates can be made based on theory, back of the envelope calculations. It is less expensive to build this system than to build an accurate simulator for the system, and the actual system will run in real time. There is also the motivational issue. It will be extremely motivating to have a prototype system ready to go.

Prior NSF Supported Work:

NSF Award IIS-1563807 (PI: Geyer, Atkeson co-PI), 8/1/16-7/31/20. $317,007 in year 1. RI: Medium: Combining Optimal and Neuromuscular Controllers for Agile and Robust Humanoid Behavior, started Aug 1, 2016, and has been active for three months. Because this award is so new, we report on a previous NSF award.

(a) NSF award: IIS-0964581 (PI: Hodgins, Atkeson co-PI); *amount:* $699,879; *period:* 7/1/10 - 6/30/14.
(b) Title: RI: Medium: Collaborative Research: Trajectory Libraries for Locomotion on Rough Terrain.
(c) Summary of Results: This grant has supported work on a variety of approaches to controlling humanoid robots based on trajectory libraries.

Intellectual Merit: The primary goal of our work on *Trajectory Libraries for Locomotion on Rough Terrain* was to develop control systems for humanoid robots that show human levels of competence, robustness and flexibility in locomotion on human-scale rough terrain, and explore library approaches to generating behavior. Results include: 1) A hierarchical approach to online optimal control of behavior. Low level behavior (accelerations, joint torques, and contact forces) is optimized on a very fast time scale (1ms) using quadratic programming. Longer term behavior (center of mass trajectory, for example) is optimized for several seconds using Differential Dynamic Programming (a 2nd order gradient based trajectory optimization technique). This work received a "Best Oral Paper Award" at Humanoids 2013. 2) Multiple model policy optimization - algorithms for training robust policies/controllers using multiple models. This approach forms the basis for new algorithms to learn and optimize policies using model-based approaches, greatly speeding up reinforcement learning and developing a new approach to robust learning. 3) Learning of optimized trajectories - simple learning approaches to concisely represent and accurately predict optimal trajectories. 4) Globally optimal control of instantaneously coupled systems (ICS), which is designed by coordinating multiple lower-dimensional optimal controllers. We augmented subsystems of the ICS with coordination variables, and then used value functions to coordinate the augmented subsystems by managing tradeoffs of the coordination variables. 5) We discovered that many features of optimized walking including costs can be fit using simple global function approximation, such as quadratic function approximation. 6) A paradigm for designing controllers for complex systems, "informed priority control", which coordinates multiple sub-policies. These tools enable us to prototype complex control system designs faster. 7) State estimation for mobile and humanoid robots with "floating body" dynamics.

We have made contributions to planning and control of human-like motion for humanoid characters and robots. We have made contributions to the control of low impedance robots and similar systems. We have made contributions to understanding human standing balance and walking. Our findings have contributed to progress in planning algorithms, and also progress in nonlinear optimal control. More specifically we have contributed to the planning and optimal control of high dimensional systems such as humanoid robots. Our work led to an

Figure 13: DARPA Robotics Challenge.

excellent performance of our team in the DARPA Robotics Challenge, where we were the only robot to try all tasks and not fall down or need to be rescued by humans (Figure 13).

Broader Impacts. We developed more useful robots and knowledge that may help with a significant social problem, why people fall and injure themselves. We coordinated our outreach activities with the larger outreach efforts of CMU's NSF Engineering Research Center on Quality of Life Technology to scale up reach and effectiveness. Our technologies are being shared by being published, and papers are available electronically. The technologies were demonstrated on entries in the DARPA Robotics Challenge. We created and taught relevant courses, including an undergraduate course on humanoid robotics and a graduate course on optimization of behavior. We put the teaching materials on the web to widely disseminate the results. Our work is being used by Disney Research and through this technology transfer will eventually be used in entertainment and education applications, and will be available to and inspire the public. A graduate student participated on a Discovery Channel TV series, "The Big Brain Theory: Pure Genius". One purpose of the TV series was getting people excited about engineering. A graduate student supported a group of all female high school students in the robot FIRST competition. A robot character in a Disney movie was inspired by our work on soft robots (Baymax in Big Hero 6) [8].

Development of Human Resources. The project involved five PhD graduate students and four postdocs. We had weekly individual meetings, weekly lab meetings, and we individually mentored all participants. All participants actively did research, made presentations to our group, gave conference presentations, and gave lectures in courses. The students served as teaching assistants.

(d) Publications resulting from this NSF award: [98, 60, 2, 48, 97, 73, 81, 78, 80, 89, 102, 77, 70, 3, 103, 7, 44, 86, 91, 31, 53, 4, 88, 79, 85, 89, 87, 43, 20]. This award led to further work supported by DARPA: [21, 22, 23, 90, 92, 12, 11, 54, 9].

(e) Other research products: We have made our motion data available in the CMU motion capture database [26].

(f) Renewed support. This proposal is not for renewed support.

Broader Impacts [of the Proposed Work]:

One theory as to why robots are so clumsy is that they have little or no tactile sensing. Endowing robots with superhuman sensing may allow them to be more useful, and cheaper to program or train. **We expect to make programming robots to do tasks involving contact much cheaper.** The proposed system is critical for **safety** of robots that are near or touch humans, and for improved robot **performance** in many manipulation tasks. Possible applications of this approach include household robots that interact with humans, aware furniture and car interiors, aware clothes, aware tools, and aware floors, walls, and ceilings. Optical skin-based sensing is a practical, manufacturable, and maintainable approach to sensing for touching and helping people, and can be applied beyond robotics to better seats and other furniture surfaces that touch humans, and more intelligent tools and control surfaces that humans touch.

Education and Outreach. A major outreach initiative led by Atkeson is the creation of a physical and virtual Robot Museum. So far we have created exhibits on juggling robots, robot actuation (gears vs. direct drive), mobile robots, soft robots, Steve Jacobsen and Sarcos, robots in literature, legged robots, computer graphics (Ivan Sutherland), and AI (Newell and Simon). Our next major initiatives are 1) to develop cell phone apps that trigger off augmented reality (AR) tags and robot pictures in halls to provide a self-guided tour of the Robotics Institute, and 2) use virtual reality (VR) to provide access to our collection from anywhere in the world. We want anyone to be able to design, build, debug, evaluate, and repair a historical robot in virtual reality. The impact of Atkeson's work will be increased by a new Disney TV show based on the characters from the Disney movie Big Hero 6, including the inflatable medical robot Baymax inspired by Atkeson's work on inflatable robots. We have coordinated our outreach activities with the larger outreach efforts of CMU's Robotics Institute to scale up reach and effectiveness. Our technologies are being shared by being published, and papers and software are available electronically.

Development of Human Resources. This project will fund a series of postdoctoral fellows. See the Postdoc Mentoring statement for more information. We expect graduate students funded from other sources and undergraduates to also participate.

Participation of Underrepresented Groups. We will make use of ongoing efforts in the Robotics Institute and CMU-wide. These efforts include supporting minority visits to CMU, recruiting at various conferences and educational institutions, and providing minority fellowships. As the Robotics Institute PhD admissions chair in 2016, Atkeson led a process which resulted in 31% of acceptances going to female applicants. As a member of the Robotics Institute hiring committee in 2016, Atkeson participated in a process that led to 10 out of 18 interviewees being female. Atkeson is assisting efforts at CMU to raise money for fellowships for students who will help us in our efforts to serve diverse populations and communities, including our own.

Dissemination Plan. For a more complete description of our dissemination plan, see our Data Management Plan. We will maintain a public website to freely share our simulations and control code, and to document research progress with video material. We will present our work at conferences and publish it in journals, and will use these vehicles to advertise our work to potential collaborators in science and industry.

Technology Transfer. Our research results and algorithms are being used by Disney Research and through this technology transfer path will eventually be used in entertainment and education applications, and will be available to and inspire the public. Two recent postdocs work at Disney Research. Three recent students work at Boston Dynamics transferring our work to industrial applications, one recent student and recent postdoc work on self-driving cars at Uber, one recent student works on self-driving cars at Apple, and one recent student works on humanoid robotics at the Toyota Research Institute. An older former student is the CTO of the Amazon drone effort. Several older former students work at Google. We are thrilled that we and our students are part of the robotics revolution.

Curriculum Development Activities:

We will develop course material on robot learning and reasoning, which will directly be influenced by the planned activities of this proposal and freely available on the web. The PIs currently teach several courses that will benefit from this material. For example, *16-745: Dynamic Optimization* and *16-711: Kinematics, Dynamics, and Control,* directly address the research areas in which this proposal is embedded. We also teach a course designed to attract undergraduates into the field, *16-264: Humanoids.* All of these courses emphasize learning from interaction by working with real robots.

References

[1] B. Ali, M.A. Ayub, and H. Yussof. Characteristics of a new optical tactile sensor for interactive robot fingers. *Intl. J. Soc. Robot*, 4:85–91, 2012.

[2] S.O. Anderson and J.K. Hodgins. Adaptive torque-based control of a humanoid robot on an unstable platform. In *Humanoid Robots (Humanoids), 2010 10th IEEE-RAS International Conference on*, pages 511–517, 2010.

[3] S.O. Anderson and J.K. Hodgins. Informed priority control for humanoids. In *Humanoid Robots (Humanoids), 2011 11th IEEE-RAS International Conference on*, pages 416–422, 2011.

[4] Stuart Anderson. *The Design of Control Architectures for Force-controlled Humanoids Performing Dynamic Tasks*. PhD thesis, Robotics Institute, Carnegie Mellon University, 2013.

[5] M. S. Asif, A. Ayremlou, A. Sankaranarayanan, A. Veeraraghavan, and R. Baraniuk. Flatcam: Thin, lensless cameras using coded aperture and computation. In *TCI*, 2016.

[6] Tareq Assaf, Calum Roke, Jonathan Rossiter, Tony Pipe, and Chris Melhuish. Seeing by touch: Evaluation of a soft biologically-inspired artificial fingertip in real-time active touch. *Sensors*, 14(2):2561, 2014.

[7] C. G. Atkeson. Efficient robust policy optimization. In *American Control Conference*, 2012.

[8] C. G. Atkeson. Big Hero 6: Let's Build Baymax. `build-baymax.org`, 2015.

[9] C. G. Atkeson, B. P. W. Babu, N. Banerjee, D. Berenson, C. P. Bove, X. Cui, M. DeDonato, R. Du, S. Feng, M. Gennert, J. P. Graff, P. He, A. Jaeger, J. Kim, K. Knoedler, L. Li, C. Liu, X. Long, T. Padir, F. Polido, G. G. Tighe, and X. Xinjilefu. NO FALLS, NO RESETS: Reliable humanoid behavior in the DARPA Robotics Challenge. In *IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, 2015.

[10] Christopher Atkeson et al. NO FALLS, NO RESETS: Reliable humanoid behavior in the DARPA Robotics Challenge. In *the 15th IEEE-RAS International Conference on Humanoid Robots (Humanoids'15)*, 2015.

[11] Benzun Pious Wisely Babu, Ruixiang Du, Taskin Padir, and Michael Gennert. Improving robustness in complex tasks for a supervisor operated humanoid˙ In *IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, 2015.

[12] Nandan Banerjee, Xianchao Long, Ruixiang Du, Felipe Polido, Siyuan Feng, Christopher G. Atkeson, Michael Gennert, and Taskin Padir. Human-supervised control of the ATLAS humanoid robot for traversing doors. In *IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, 2015.

[13] S. Begej. Planar and finger-shaped optical tactile sensors for robotic applications. *IEEE Journal on Robotics and Automation*, 4(5):47200484, 1988.

[14] M. Blank and J.F. Antaki. Low-cost quantitative palpation device for breast lesion tracking and telehealth. In *IEEE Health Innovations and Point-of-care Technologies Conference*, pages 280–283, 2014.

[15] Bill Chiu, Eamonn Keogh, and Stefano Lonardi. Probabilistic discovery of time series motifs. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 493–498. ACM, 2003.

[16] C. Chorley, C. Melhuish, T. Pipe, and J. Rossiter. Development of a tactile sensor based on biologically inspired edge encoding. In *Advanced Robotics, 2009. ICAR 2009. International Conference on*, pages 1–6, 2009.

[17] M. Cutkowsky and W. Provanchar. Force and torque sensing. In *Springer Handbook of Robotics*, pages 717–736, 2016.

[18] R. S. Dahiya and M. Valle. *Robotic Tactile Sensing: Technologies and System*. Springer, 2013.

[19] DLR. Artificial skin - a robot's personal touch. `http://www.dlr.de/rm/en/desktopdefault.aspx/tabid-3816/6236_read-9002/`, 2016.

[20] S. Feng, X. Xinjilefu, W. Huang, and C. G. Atkeson. 3d walking based on online optimization. In *IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, 2013.

[21] Siyuan Feng, Eric Whitman, X Xinjilefu, and Christopher G. Atkeson. Optimization Based Full Body Control for the Atlas Robot. In *Proc IEEE Conf Humanoids*, Madrid, Spain, 2014.

[22] Siyuan Feng, Eric Whitman, X Xinjilefu, and Christopher G. Atkeson. Optimization-based full body control for the DARPA Robotics Challenge. *Journal of Field Robotics*, 32(2):293–312, 2015.

[23] Siyuan Feng, X Xinjilefu, Christopher G. Atkeson, and Joohyung Kim. Optimization based controller design and implementation for the Atlas robot in the DARPA Robotics Challenge Finals. In *IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, 2015.

[24] FLIR. The power of micro thermal imaging. `www.flir.com/cores/lepton`, 2016.

[25] Google. Project soli. `https://atap.google.com/soli/`, 2016.

[26] J Hodgins. mocap.cs.cmu.edu.

[27] K. Hoshino, D. Mori, and M. Tomida. An optical tactile sensor assuming cubic polynomial deformation of elastic body. *Journal of Robotics and Mechatronics*, 21(6):780–788, 2009.

[28] R. Howe. Harvard biorobotics lab. `http://biorobotics.harvard.edu/research.html`, 2016.

[29] D. Hristu, N. Ferrier, and R. W. Brockett. The performance of a deformable-membrane tactile sensor: basic results on geometrically-defined tasks. In *Robotics and Automation, 2000. Proceedings. ICRA '00. IEEE International Conference on*, volume 1, pages 508–513, 2000.

[30] D. Hristu, N. Ferrier, and R.W. Brockett. The performance of a deformable-membrane tactile sensor: basic results on geometrically-defined tasks. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 508–513, 2000.

[31] Weiwei Huang, Junggon Kim, and Christopher G. Atkeson. Energy-based optimal step planning for humanoids. In *Robotics and Automation (ICRA), 2013 IEEE International Conference on*, pages 3124–3129, 2013.

[32] D. Hughes and N. Correll. Texture recognition and localization in amorphous robotic skin. *Bioinspiration & Biomimetics*, 10(5):1–20, 2015.

[33] N. Igo and K. Hoshino. Small optical tactile sensor for robots. In *IEEE/SICE International Symposium on System Integration (SII)*, pages 746–751, 2012.

[34] Takuya Ikai, Shota Kamiya, and Masahiro Ohka. Robot control using natural instructions via visual and tactile sensations. *Journal of Computer Science*, 12(5):246–254, 2016.

[35] Y. Ito, Y. Kim, C. Nagai, and G. Obinata. Shape sensing by vision-based tactile sensor for dexterous handling of robot hands. In *2010 IEEE International Conference on Automation Science and Engineering*, pages 574–579, 2010.

[36] Y. Ito, Y. Kim, and G. Obinata. Robust slippage degree estimation based on reference update of vision-based tactile sensor. *IEEE Sensors Journal*, 11(9):2037–2047, 2011.

[37] Y. Ito, Y. Kim, and G. Obinata. Contact region estimation based on a vision-based tactile sensor using a deformable touchpad. *Sensors*, 14:5805–5822, 2014.

[38] A.R. Johnston. Optical proximity sensors for manipulators, 1973. NASA Technical Memorandum 33-612.

[39] K. Kamiyama, H. Kajimoto, N. Kawakami, and S. Tachi. Evaluation of a vision-based tactile sensor. In *Robotics and Automation, 2004. Proceedings. ICRA '04. 2004 IEEE International Conference on*, volume 2, pages 1542–1547, 2004.

[40] K. Kamiyama, H. Kajimoto, N. Kawakami, and S. Tachi. Evaluation of a vision-based tactile sensor. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 1542–1547, 2004.

[41] K. Kamiyama, K. Vlack, T. Mizota, H. Kajimoto, N. Kawakami, and S. Tachi. Vision-based sensor for real-time measuring of surface traction fields. *IEEE Computer Graphics and Applications*, 25(1):68–75, 2005.

[42] Jaakko Karras. Tactile sensors for palm-size crawling robots. Master's thesis, EECS Department, University of California, Berkeley, May 2014.

[43] J. Kim, N. S. Pollard, and C. G. Atkeson. Quadratic encoding of optimized humanoid walking. In *IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, 2013.

[44] S. Kim, C. G. Atkeson, and S. Park. Perturbation-dependent selection of postural feedback gain and its scaling. *Journal of Biomechanics*, 45(8):1379–1386, 2012.

[45] E. Knoop and J. Rossiter. Dual-mode compliant optical tactile sensor. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 1006–1011, 2013.

[46] S. J. Koppal, S. Yamazaki, and S. G. Narasimhan. Exploiting dlp illumination dithering for reconstruction and photography of high-speed scenes. *In Press, IJCV*, 2011.

[47] K. Kuhnlenz, B. Kuhnlenz, R. Dahiya, and G. Cheng. Workshop on novel trends in multimodal robotic skins, 2014. http://mysite.hs-coburg.de/ kuko140m/mros2014.htm.

[48] T. Kwon and J. K. Hodgins. Control systems for human running using an inverted pendulum model and a reference motion capture sequence. In *Proceedings of the 2010 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, SCA '10, 2010.

[49] N. F. Lepora and B. Ward-Cherrier. Superresolution with an optical tactile sensor. In *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*, pages 2686–2691, 2015.

[50] N.F. Lepora and B. Ward-Cherrier. Superresolution with an optical tactile sensor. In *IEEE Intl. Conf. on Intelligent Robots and Systems (IROS)*, pages 2686–2691, 2015.

[51] R. Li, R. Platt, W. Yuan, A. ten Pas, N. Roscup, M. A. Srinivasan, and E. Adelson. Localization and manipulation of small parts using gelsight tactile sensing. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3988–3993, 2014.

[52] Jessica Lin, Eamonn Keogh, Li Wei, and Stefano Lonardi. Experiencing sax: a novel symbolic representation of time series. *Data Mining and Knowledge Discovery*, 15(2):107–144, 2007.

[53] C. Liu, C. G. Atkeson, and J. Su. Biped walking control using a trajectory library. *ROBOTICA*, 31:311–322, 2013.

[54] Chenggang Liu, Christopher G Atkeson, Siyuan Feng, and X Xinjilefũ. Full-body motion planning and control for the car egress task of the DARPA Robotics Challenge. In *IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, 2015.

[55] Luxexcel. Home page. `www.luxexcel.com`, 2016.

[56] Marianna Madry, Liefeng Bo, Danica Kragic, and Dieter Fox. St-hmp: Unsupervised spatio-temporal feature learning for tactile data. In *IEEE International Conference on Robotics and Automation (ICRA)(to appear)*, 2014.

[57] H. Maekawa, K. Tanie, K. Komoriya, M. Kaneko, C. Horiguchi, and T. Sugawara. Development of a finger-shaped tactile sensor and its evaluation by active touch. In *Robotics and Automation, 1992. Proceedings., 1992 IEEE International Conference on*, volume 2, pages 1327–1334, 1992.

[58] U. Martinez-Hernandez. Tactile sensors. `http://www.scholarpedia.org/article/Tactile_Sensors`, 2015.

[59] Colin McManus, Ben Upcroft, and Paul Newmann. Scene signatures: localised and point-less features for localisation. *Robotics: Science and Systems*, 2014.

[60] M. Mistry, A. Murai, K. Yamane, and J. Hodgins. Sit-to-stand task on a humanoid robot from human demonstration. In *Humanoid Robots (Humanoids), 2010 10th IEEE-RAS International Conference on*, pages 218–223, 2010.

[61] K. Nagata, M. Ooki, and M. Kakikur. Feature detection with an image based compliant tactile sensor. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 838–843, 1999.

[62] University of Bielefeld Neuroinformatics Group. Humanoids 2016 workshop. `https://ni.www.techfak.uni-bielefeld.de/ICHR2016WS/home`, 2016.

[63] K. Noda, H. Arie, Y. Suga, and T. Ogata. Multimodal integration learning of robot behavior using deep neural networks. *Robotics and Autonomous Systems*, 62(6):721–736, 2014.

[64] M. Ohka, Y. Mitsuya, K. Hattori, and I. Higashioka. Data conversion capability of optical tactile sensor featuring an array of pyramidal projections. In *IEEE Intl. Conf. on Multisensor Fusion and Integration for Intelligent Systems*, 1996.

[65] Optoforce. 3D Force Sensor, 2014. www.optoforce.com.

[66] Radhen Patel and Nikolaus Correll. Integrated force and distance sensing using elastomer-embedded commodity proximity sensors. In *Proceedings of Robotics: Science and Systems*, 2016.

[67] Rethink Robotics. Using the camera, 2016. `github.com/RethinkRobotics/sdk-docs/wiki/Using-the-Cameras`.

[68] Robotiq. Cameras, 2016. `robotiq.com/products/camera/`.

[69] S. Saga, S Kuroki, N. Kawakami, and S. Tachi. Fibratus tactile sensor using reflection on an optical lever. In *Siggraph*, 2007.

[70] S. Sanan, M. H. Ornstein, and C. G. Atkeson. Physical human interaction for an inflatable manipulator. In *IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 7401–7404, 2011.

[71] K. Sato, K. Kamiyama, N. Kawakami, and S. Tachi. Finger-shaped GelForce: sensor for measuring surface traction fields for robotic hand. *IEEE Transactions on Haptics*, 3(1):37–47, 2010.

[72] K. Sato, K. Kamiyama, H. Nii, N. Kawakami, and S. Tachi. Measurement of force vector field of robotic finger using vision-based haptic sensor. In *IEEE Intl. Conf. on Intelligent Robots and Systems (IROS)*, pages 488–493, 2008.

[73] S. Schaal and C. G. Atkeson. Learning control for robotics. *IEEE Robotics & Automation Magazine*, 17(2):20–29, 2010.

[74] J. L. Schneiter and T. B. Sheridan. An optical tactile sensor for manipulators. *Robotics & Computer Integrated Manufacturing*, 1(1):65–71, 1984.

[75] H. Shinoda and S. Ando. Ultrasonic emission tactile sensor for contact localization and characterization. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 2536–2543, 1994.

[76] D.C. Sims, Y. Yue, and S.K. Nayar. Towards flexible sheet cameras: Deformable lens arrays with intrinsic optical adaptation. In *IEEE International Conference on Computational Photography (ICCP)*, 2016.

[77] Kwang Won Sok, Katsu Yamane, Jehee Lee, and Jessica Hodgins. Editing dynamic human motions via momentum and force. In *Proceedings of the 2010 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, SCA '10, pages 11–20, 2010.

[78] B. Stephens. Dynamic balance force control for compliant humanoid robots. In *International Conference on Intelligent Robots and Systems (IROS)*, pages 1248–1255, 2010.

[79] B. Stephens. *Push Recovery Control for Force-Controlled Humanoid Robots*. PhD thesis, Carnegie Mellon University, 2011.

[80] Benjamin J. Stephens and Christopher G. Atkeson. Push Recovery by stepping for humanoid robots with force controlled joints. *2010 10th IEEE-RAS Int. Conf. Humanoid Robot.*, pages 52–59, 2010.

[81] Martin Stolle and Christopher G. Atkeson. Finding and transferring policies using stored behaviors. *Autonomous Robots*, 29(2):169—200, 2010.

[82] J. Ueda, Y. Ishida, M. Kondo, and T. Ogasawara. Development of the NAIST-hand with vision-based tactile fingertip sensor. In *???*, 2005.

[83] Jun Ueda, Y. Ishida, M. Kondo, and T. Ogasawara. Development of the NAIST-Hand with vision-based tactile fingertip sensor. In *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, pages 2332–2337, 2005.

[84] D. Um, B. Stankovic, K. Giles, T. Hammond, and V. Lumelsky. A modularized sensitive skin for motion planning in uncertain environments. In *IEEE International Conference on Robotics and Automation (ICRA)*, volume 1, pages 7–12, 1998.

[85] E. Whitman and C. G. Atkeson. Control of instantaneously coupled systems applied to humanoid walking. In *International Conference on Humanoid Robots*, pages 210–217, 2010.

[86] E. C. Whitman and C. G. Atkeson. Multiple model robust dynamic programming. In *American Control Conference (ACC)*, pages 5998–6004, 2012.

[87] E. C. Whitman, B. J. Stephens, and C. G. Atkeson. Torso rotation for push recovery using a simple change of variables. In *International Conference on Humanoid Robots*, pages 50–56, 2012.

[88] Eric Whitman. *Coordination of Multiple Dynamic Programming Policies for Control of Bipedal Walking*. PhD thesis, Robotics Institute, Carnegie Mellon University, 2013.

[89] D. Xing, C. G. Atkeson, J. Su, and B. Stephens. Gain scheduled control of perturbed standing balance. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4063–4068, 2010.

[90] Xinjilefu. *State Estimation for Humanoid Robots*. PhD thesis, Robotics Institute, Carnegie Mellon University, 2015.

[91] Xinjilefu and C.G. Atkeson. State estimation of a walking humanoid robot. In *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, pages 3693–3699, 2012.

[92] X Xinjilefu, Siyuan Feng, and Christopher G. Atkeson. Center of mass estimator for humanoids and its application in modelling error compensation, fall detection and prevention. In *IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, 2015.

[93] Y. Yamada, T. Morizono, Y. Umetani, and H. Takahashi. Highly soft viscoelastic robot skin with a contact object-location-sensing capability. *IEEE Transactions on Industrial Electronics*, 52(4):960–968, 2005.

[94] Y. Yamada, I. Yoshihiro, M. Fukunaga, N. Fujimoto, E. Ohta, T. Morizono, and Y. Umetani. Soft viscoelastic robot skin capable of accurately sensing contact location of objects. In *IEEE Intl. Conf. on Multisensor Fusion and Integration for Intelligent Systems*, pages 105–110, 1999.

[95] A. Yamaguchi and C. Atkeson. Finger vision video. `https://youtu.be/ifOwQdy9gDg`, 2016.

[96] Akihiko Yamaguchi and Christopher G. Atkeson. Combining finger vision and optical tactile sensing: Reducing and handling errors while cutting vegetables. In *the 16th IEEE-RAS International Conference on Humanoid Robots (Humanoids'16)*, 2016. (In press. Author's version is here: `http://akihikoy.net/p/hm16b.pdf` video: `https://youtu.be/ifOwQdy9gDg`).

[97] K. Yamane, S.O. Anderson, and J.K. Hodgins. Controlling humanoid robots with human motion data: Experimental validation. In *Humanoid Robots (Humanoids), 2010 10th IEEE-RAS International Conference on*, pages 504–510, 2010.

[98] K. Yamane and J. Hodgins. Control-aware mapping of human motion data with stepping for humanoid robots. In *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pages 726–733, 2010.

[99] Lexiang Ye and Eamonn Keogh. Time series shapelets: a new primitive for data mining. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 947–956. ACM, 2009.

[100] W. Yuan, R. Li, M.A. Srinivasan, and E.H. Adelson. Measurement of shear and slip with a gelsight tactile sensor. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 304–311, 2015.

[101] Hanafiah Yussof, Jiro Wada, and Masahiro Ohka. Sensorization of robotic hand using optical three-axis tactile sensor: Evaluation with grasping and twisting motions. *Journal of Computer Science*, 6(8):955–962, 2010.

[102] M. Zucker, J. A. Bagnell, C. G. Atkeson, and J. Kuffner. An optimization approach to rough terrain locomotion. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 3589–3595, 2010.

[103] M. Zucker, N. D. Ratliff, M. Stolle, J. E. Chestnutt, J. A. Bagnell, C. G. Atkeson, and J. Kuffner. Optimization and learning for rough terrain legged locomotion. *International Journal of Robotic Research*, 30(2):175–191, 2011.