GravitySpace: Tracking Users and Their Poses in a Smart Room Using a Pressure-Sensing Floor

Alan Bränzel, Christian Holz, Daniel Hoffmann, Dominik Schmidt, Marius Knaust,
Patrick Lühne, René Meusel, Stephan Richter, Patrick Baudisch
Hasso Plattner Institute, Potsdam, Germany
{firstname.lastname}@student.hpi.uni-potsdam.de
{christian.holz, dominik.schmidt, patrick.baudisch}@hpi.uni-potsdam.de

ABSTRACT
We explore how to track people and furniture based on a high-resolution pressure-sensitive floor. Gravity pushes people and objects against the floor, causing them to leave imprints of pressure distributions across the surface. While the sensor is limited to sensing direct contact with the surface, we can sometimes *conclude* what takes place *above* the surface, such as users' poses or collisions with virtual objects. We demonstrate how to extend the range of this approach by sensing through passive furniture that propagates pressure to the floor. To explore our approach, we have created an 8 m^2 back-projected floor prototype, termed *GravitySpace*, a set of passive touch-sensitive furniture, as well as algorithms for identifying users, furniture, and poses. Pressure-based sensing on the floor offers four potential benefits over camera-based solutions: (1) it provides consistent coverage of rooms wall-to-wall, (2) is less susceptible to occlusion between users, (3) allows for the use of simpler recognition algorithms, and (4) intrudes less on users' privacy.

Author Keywords
Interactive floor; smart rooms; ubicomp; multitoe; multi-touch; FTIR; tabletop; vision.

ACM Classification Keywords
H.5.2. [Information Interfaces and Presentation]: User Interfaces–Input devices and strategies, interaction styles.

General Terms
Design; Human Factors.

INTRODUCTION
Brummit et al. define self-aware spaces as "… [a space that] knows its own geometry, the people within it, their actions & preferences, and the resources available to satisfy their requests" [5]. Such smart rooms support users by offering not only a series of convenient functions, like home automation, but also by acting pro-actively on the user's behalf. Similar systems have been proposed to monitor the wellbeing of (elderly) inhabitants [15].

In order to provide this support, smart rooms track users and try to automatically recognize their activities. In systems like EasyLiving, this was done by pointing tracking equipment, such as cameras, at the interior of the room [5]. The direct observation of scenes using computer vision is of limited reliability, because of illumination and perspective effects, as well as occlusion between people. The latter also affects more recent approaches based on depth cameras (e.g., *LightSpace* [43]).

We propose an alternative approach to tracking people and objects in smart rooms. Building on recent work on touch-sensitive floors (e.g., *Multitoe* [1]) and pose reconstruction, such as [44], we explore how much a room can infer about its inhabitants solely based on the pressure imprints people and objects leave on the floor.
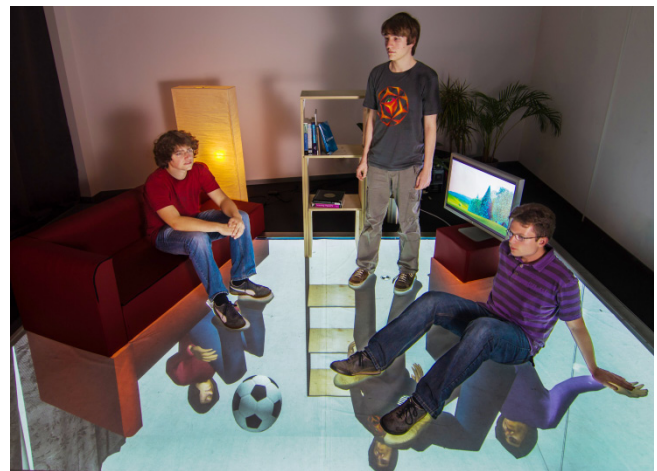

Figure 1: GravitySpace recognizes people and objects. We use a mirror-metaphor to show how GravitySpace identifies users and tracks their location and poses, solely based on the pressure imprints they leave on the floor.

GRAVITYSPACE
Figure 1 shows our floor installation *GravitySpace* with three users and three pieces of furniture. To illustrate what the system senses and reconstructs about the physical world, the prototype displays its understanding of the physical world using a mirror-metaphor, so that every object stands on its own virtual reflection. Based on this mirror world, we see that GravitySpace recognizes the position and orientation of multiple users, the identity of users as demonstrated by showing their personalized avatars, selected poses, such as stand-

ing and sitting on the floor and on furniture, and tracking of leg movements to interact with virtual objects, here a soccer ball. GravitySpace updates in real-time and runs a physics engine to model the room above the surface. To convey the 3D nature of the sensing, this photo was shot with a tracked camera—this camera tracking is *not* part of GravitySpace.

GravitySpace consists of a single sensor, namely the floor itself which is pressure-sensitive, while the seating furniture passively propagates pressure to the floor. All the tracking and identification shown in Figure 1 is solely based on pressure imprints objects leave on this floor as shown in Figure 3.
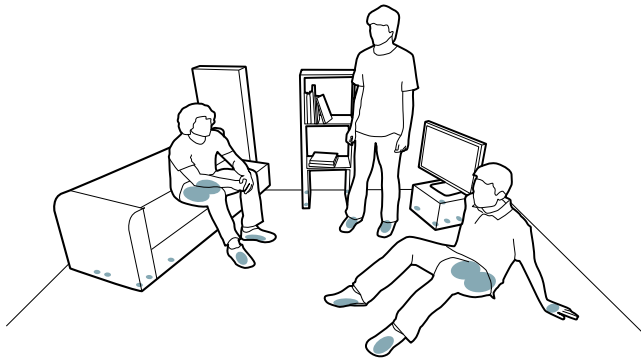
Figure 2: Gravity pushes people and objects against the ground, where they leave *imprints* that GravitySpace can sense.

Our approach is based on the general principle of gravity, which pushes people and objects against the floor, causing the floor to sense pressure imprints as illustrated in Figure 3. While the pressure sensor is limited to sensing *contact* with the ground, GravitySpace not only tracks what happens in the floor plane (such as shoeprints), but is able to draw a certain amount of *conclusions* about what happens in the space *above* the floor, such as a user's pose, or the collision between a user and a virtual ball. In addition, GravitySpace senses what takes place on top of special furniture that propagates pressure to the floor.
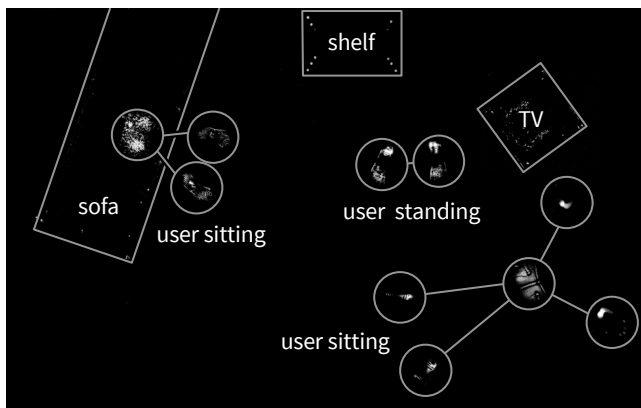
Figure 3: GravitySpace sees the scene from Figure 1 as a set of imprints (circles, lines, and text added for clarity).

Figure 3 shows the scene from Figure 1 as perceived by GravitySpace. This is the information GravitySpace uses to reconstruct the scene above the ground. We see four concepts here, which are detailed in Section *Algorithms*:

(1) recognition of poses (86.12% accuracy) based on classifying contact types, such as hands or buttocks, and their spatial arrangement, (2) prediction of leg movements by analyzing pressure distributions, and (3) pressure-based markers that allow GravitySpace to detect objects, such as furniture. In addition, GravitySpace recognizes users based on their shoeprints, similar to *Multitoe* [1], but optimized for the 20 times larger floor size and a larger number of simultaneous users (99.82% accuracy with 120 users in real time).

Prototype Hardware

Figure 4 shows our current GravitySpace prototype hardware. It senses pressure based on FTIR [13], using a camera located below the floor. It provides an 8 m^2 interaction surface in a single seamless piece and delivers 12 megapixels overall pressure sensing resolution at a pixel size of 1×1 mm. Our prototype also offers 12 megapixel back projection. While not necessary for tracking, it allows us to visualize the workings of the system, as we did in Figure 1.
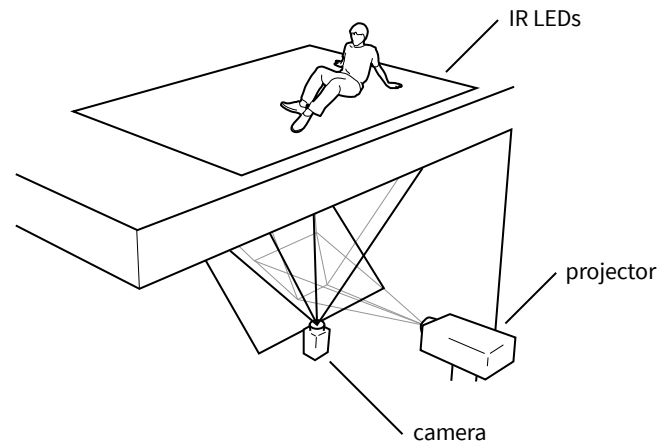
Figure 4: The GravitySpace prototype senses 25 dpi pressure and projects across an active area of 8 m² in a single seamless piece.

We expect sensing hardware of comparable size and resolution to soon be inexpensive and mass available, for example in the form of a large, thin, high-resolution pressure sensing foil (e.g., *UnMousePad* [34]). We envision this material to be integrated into carpet and as such installed in new homes wall-to-wall. Since the technology is not quite ready to deliver the tens of megapixel resolution we require for an entire room, our FTIR-based prototype allows us to explore our vision of tracking based on pressure imprints *today*.

CONTRIBUTION

The main contribution of this paper is a new approach to tracking people and objects in a smart room, namely based on a high-resolution pressure-sensitive floor. While the sensor is limited to sensing *contact* with the surface, we demonstrate how to conclude a range of objects and events that take place *above* the surface, such as a user's pose and collisions with virtual objects. We demonstrate how to extend the range of this approach by sensing *through* passive furniture that propagates pressure to the floor.

To explore this approach, we have designed and implemented a fully-functional room-size prototype, which we use to

demonstrate our vision true to scale. We have also implemented algorithms for tracking and identifying users, furniture, and poses.

Our main goal with this paper is to demonstrate the technology. We validate it using a technical evaluation of user identification and pose recognition accuracy. The philosophy behind smart rooms, their applications and usability, however, are outside the scope of this paper.

Benefits and Limitations

Compared to traditional camera-based solutions, the proposed approach offers four benefits:

(1) It provides consistent coverage of rooms, wall-to-wall. Camera-based systems have a pyramid-shaped viewing space. Motion capture installations resolve this by leaving space along the edges, but that is impractical in regular rooms, leading to uneven or spotty coverage. Floor-based tracking in contrast, can be flat, integrated into the room itself, and provides consistent coverage across the room.

(2) It is less susceptible to occlusion between users. The perspective from below is particularly hard to block—simply because people tend to stand next to each other, resulting in discernible areas of contact. From a more general perspective, the benefit of pressure sensing is that mass is hard to "hide". Mass has to manifest itself somewhere, either through direct contact or indirectly through another object it is resting on. Camera-based systems, in contrast, may suffer from users occluding each other if the cameras mounted in one spot (e.g., [43]). Systems distributing multiple cameras around a room (e.g., [21]) still suffer from dead spots (e.g., in the midst of groups of users).

(3) *It allows for the use of simpler, more reliable recognition algorithms.* Our approach reduces the recognition problem from comparing 3D objects to comparing flat objects, because all objects are flat when pressed against a rigid surface. This limits objects to three degrees of freedom (translation and rotation in the plane) and allows us to match objects using simple, robust, and well-understood algorithms from digital image processing [8, 11].

(4) *Pressure-based tracking is less privacy-critical.* While floor-based sensing captures a lot of information relevant to assisted living applications (e.g., [15]), it never captures photos or video of its inhabitants, mitigating privacy concerns (e.g., while getting dressed or using the bathroom).

On the other hand, our floor-based approach is obviously limited in that it can recognize objects only when they are in direct contact with the floor. While we reduce the impact of these limitations using 3D models based on inverse kinematics, events taking place in mid-air can obviously not be sensed, such as the angle of an arm being raised or a user's gaze direction. The approach is also inherently subject to lag in that the floor learns about certain events only with a delay. We cannot know the exact position of a user sitting down until the user makes contact with the seat. As we place the avatar in between, it is subject to inaccuracy.

RELATED WORK

The work presented in this paper builds on smart rooms, interactive floors, and user identification in ubiquitous computing.

Multi Display Environments and Smart Rooms

The concept of integrating computing into the environment goes back as far as Weiser (*Ubiquitous computing* [42]). The concept has been researched in the form of smart components in a room, e.g., in multi-display environments such as the Stanford *iRoom* [4] or roomware (*iLand* [37]). Alternatively, researchers have instrumented the room itself, e.g., using cameras and microphones (e.g., *EasyLiving* [5]), making user tracking a key component of the system. The *Georgia Tech Aware Home* [15] tracks users based on multi-user head tracking and combined audio and video sensing. Most recently, Wilson and Benko demonstrated how to instrument rooms using multiple depth cameras (*Lightspace* [43]).

Pressure-Sensing Floors

A series of floor prototypes have used a range of pressure sensing technologies offering a variety of resolutions. The projection-less *magic carpet* senses pressure using piezoelectric wires and a pair of Doppler radars [27]. *Z-tiles* improved on this by introducing a modular system of interlocking tiles [32]. Pressure sensing has been implemented using force-sensing resistors [41]. FootSee [44] matches foot data from a 1.6 m² pressure pad to pre-recorded pose animations of a single user with in a fixed orientation.

In the desktop world, the *UnMousePad* improves on resistive pressure sensing by reducing the number of required wire connections [34]. Based on this, Srinivasan et al. built larger-scale installations, combined with marker-based motion systems, as well as audio and video tracking [36].

Since none of the existing technologies scale to the megapixel range of resolution, GravitySpace is built on an extension of the *Multitoe* floor. *Multitoe* uses high-resolution FTIR sensing and allows users to interact using direct manipulation [1]. Other instrumented floors are tracked using ceiling-mounted cameras (e.g., *iFloor* [16]) or *front diffuse illumination* (*IGameFloor* [12]). Purposes of interactive floors include immersion (also as part of *CAVEs* [6, 17]), gaming [12], and multi-user collaborative applications [16].

Furniture that Senses Pressure

A series of research and products use pressure-sensitive devices and furniture to monitor health, for instance, to prevent Decubiti (e.g., [39]), orthopedic use inside of shoes (e.g., [10]), and to sense pose while sitting (e.g., [23, 24]).

Sensing through Objects

The pressure-transmitting furniture presented in this paper builds on the concept of sensing *through* an object. The concept has been explored in the context of tangible objects. Mechanisms include the propagation of light through holes (*stackable markers* [3]) and optical fiber (*Lumino* [2]) and the propagation of magnetic forces, sensed using pressure sensors (*Geckos* [18]).

Matching Objects in 2D

The presented work essentially reduces a 3D problem to a 2D problem, allowing us to apply well-explored traditional algorithms from *digital image processing* [8, 11].

Identifying Users

The majority of large-scale touch technologies, such as *diffused illumination* (*DI* [20]), *front-DI* [12], and *FTIR* [13] are ignorant of who touches. *DiamondTouch* improved on this by mapping users to seat positions [9]. Bootstrapper identifies tabletop users by observing the top of their shoes [33]. Recognizing fingerprints has been envisioned to identify users of touch systems [38]; Holz and Baudisch implemented this by turning a fingerprint scanner into a touch device [14]. Schmidt et al. identify tabletop users by the size and shape of their hands [35].

User identification has been used for a variety of applications including the management of access privileges [22] and to help children with Asperger syndrome learn social protocol [30]. Olwal and Wilson used RFID tags to identify objects on the table [25]. The screen-less *Smart Floor* identifies users by observing the forces and timing of the individual phases of walking [26]. While floors so far did not have enough resolution to distinguish soles, footprints have been analyzed as evidence in crime scene investigation [28]. Sole imprints and sole wear has been used to match people either by hand and using semi-automatic techniques based on local feature extractors, such as *MSER* [28, 29]. *Multitoe* distinguishes users based on their shoeprints using template-matching [1].

WALKTHROUGH

In the following walkthrough, we revisit the elements from Figure 1 including user identification and tracking, pose detection, and pressure-transmitting furniture. Given that this paper is about a new approach to tracking, this walkthrough is intended to illustrate GravitySpace's tracking capabilities; it is not trying to suggest a specific real-world application scenario. As before, we render GravitySpace's understanding of users and furniture as a virtual 3D world under the floor. A detailed description of the shown concepts and their implementation can be found in Section *Algorithms*.

Figure 5: GravitySpace detects Daniel sitting on the sofa, and identifies and tracks René walking across the floor.

Figure 5 shows Daniel on the left as he is relaxing on the sofa and GravitySpace's interpretation of the scene in the mirror world. The same scene is shown from a pressure-sensing perspective in Figure 6. GravitySpace parses this pressure image to identify the sofa based on embedded pressure markers and to locate *someone* sitting on top of the sofa based on the pressure imprint of Daniel's buttocks. The sofa is filled with elements that transmit pressure to the floor. GravitySpace combines buttocks and the two feet next to the sofa into a pose. It also identifies Daniel based on his shoeprints. Using this information, Daniel's avatar is selected from a user library and positioned onto the imprints of feet and buttocks.

Figure 6: The scene from Figure 5 as seen by the system.

When Daniel's friend, René, comes in he is likewise identified. GravitySpace positions his personalized avatar by fitting a skeleton to René's pressure imprints using inverse kinematics, based on three control points: the two imprints of René's feet in their respective orientation, as well as René's center of gravity. The latter it placed above the center of pressure between René's shoes.

As René walks across the room, GravitySpace continuously tracks his position. By observing the pressure distribution of the foot on the ground, GravitySpace predicts where the foot currently located in midair is expected to come down. This allows it to animate the avatar without having to wait for the foot to touch down.

Figure 7: Daniel and René are controlling a video game by simply leaning over, tracked by observing varying centers of pressure. GravitySpace also detects that Andreas sits down, based on the body parts in contact with the floor.

René and Daniel decide to play a video game. As shown in Figure 7, they interact with the game without a dedicated controller by simply shifting their weight to control virtual racecars. René and Daniel accelerate and brake by leaning forward or backward; they steer by leaning left and right. GravitySpace observes this through the sofa and the cube seat. Also shown in Figure 7, Andreas, a common friend has sat down on the floor to watch the other two playing. GravitySpace determines his pose based on the texture and spatial arrangement of contact points on the floor.

Figure 8: GravitySpace allows for interaction with virtual objects as leg movement is tracked above the floor by analyzing the center of pressure of the other foot.

The three friends then decide to play a round of virtual soccer (Figure 8). The game requires them to kick a virtual ball. GravitySpace cannot directly observe what happens *above* the surface. Instead it observes weight shifts within the *other* foot, and *concludes* where the kicking foot must be located. Using inverse kinematics, it places the avatars and GravitySpace's physics engine computes how the avatar kicks the ball.

PRESSURE-TRANSMITTING FURNITURE

While it is quite possible to create furniture with active pressure sensing (e.g., [24]), we have created *passive* furniture that transmits high-resolution pressure, rather than sensing it. This offloads sensing to a single centralized active sensing component, in our case the floor. Passive furniture also reduces complexity and cost, while the absence of batteries and wires makes them easy to maintain [2].

Everyday furniture already transmits pressure. Furniture imprints, however, are limited to representing overall weight and balance. While locating the center of gravity has been demonstrated by many earlier systems (e.g., VoodooIO [40] or commercial systems, such as the Wii Balance board), this limits our ability to detect activities taking place on top of the furniture (e.g., sitting on a sofa).

In order to recognize identity and poses of the object on top in more detail, we have created the furniture pieces featured in Figure 1 and the walkthrough. They transmit pressure in comparably high resolution. We accomplish this by using an array of "transmitters". Transmitters have to offer sufficient stiffness to transmit pressure (and also to support the weight

of the person or object on top). Figure 9 shows how we constructed a cube seat. We use regular drinking straws as transmitters, making the furniture light and sturdy. 1,200 straws (8 mm in diameter) fill each cube seat; 10,000 fill the sofa, which is based on the same principle. Straws are inexpensive (e.g., 80 EUR for filling the sofa). The backrest and armrest of the sofa are pressure-sensitive as well—they are filled with longer "sangria" drinking straws. We obtain the desired curved shape by cutting the straws to length a layer at a time using a laser cutter.

Figure 9: Each cube seat is filled with 1,200 drinking straws. Here one of the steel rods that form the markers is inserted.

Straws are held together by a frame made from 15 mm fiberboard. We stabilize the straws in the box using a grid made from plywood connected to the frame, which essentially subdivides the box into 3×3 independent cells. The grid minimizes skewing, thus prevents the box from tipping over. We cover the bottom of the box with Tyvek, a material that crinkles but does not stretch, which prevents the bottom from sagging, yet transmits pressure. In addition to the leather we added a thin layer of foam as cushioning to the top of the cube seats for added comfort.

Weight shifts on top of the box can cause the box to "ride up" on the straws, which can cause an edge of the box to lose traction with the ground. To assure reliable detection, we create markers from weight rods that slide freely in plastic tubes and held in by the Tyvek. We use an asymmetric arrangement of rods to give a unique ID to each piece.

ALGORITHMS

Figure 10 summarizes the pipeline we implemented to process touches that occur on our floor, including recognizing, classifying and tracking events, users and objects based on the pressure patterns they leave. We optimized our pipeline to process the entire 12 megapixel image in real-time (25 fps).
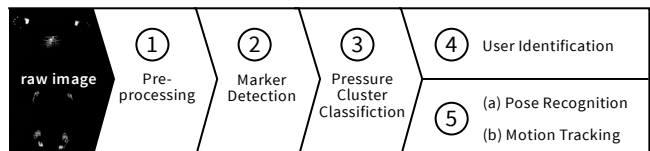

Figure 10: GravitySpace processes all input using this pipeline to recognize and track users and objects.

GravitySpace recognizes objects with texture (e.g., body parts or shoeprints) by extracting the imprint features they leave in the raw image. For objects with little discernible texture or changing texture (e.g., due to users sitting on furniture), we add features using pressure-based markers.

Our GravitySpace implementation supports three key elements: (1) Pose reconstruction using pressure cluster classification, (2) joint estimation based on pressure distributions and inverse kinematic, and (3) user identification based on shoeprints.

Step 1: Pre-Processing Pressure Images

All processing starts by thresholding the pressure image to remove noise. Our algorithm then segments this image and extracts continuous areas of pressure using a connected component analysis. In the next step, GravitySpace merges areas within close range, prioritizing areas that expand towards each other. We call the result *pressure clusters*. A pressure cluster may be for example a shoeprint or the buttocks of a sitting user. GravitySpace then tracks these clusters over time.

Step 2: Identifying Furniture Based on Markers

Pressure imprints of larger objects, such as furniture, provide little distinguishable texture on their own. In addition, the overall texture of seating furniture changes substantially when users sit down. GravitySpace therefore uses dot-based pressure-markers to locate and identify furniture. Figure 11 shows the imprint of a cube that we equipped with a marker. This particular marker consists of five points. Marker points are arranged in a unique spatial pattern that is rotation-invariant. We designed and implemented marker patterns for a sofa, several sitting cubes, and shelves.
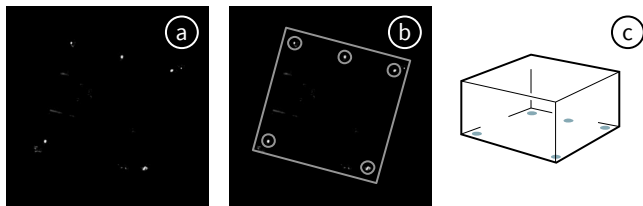

Figure 11: (a) Pressure imprint and (b) detected marker points (c) of marker-equipped cube.

To recognize markers, GravitySpace implements brute-force matching on the locations that have been classified as marker points, trying to fit each registered piece of furniture into the observed point set and minimizing the error distance. To increase the stability of recognition, our implementation keeps objects whose marker patterns have been detected in a history list and increases the confidence level of recently recognized objects. We also use hysteresis to decide when marker detection is stable based on the completeness of markers and their history confidence.

Step 3: Classifying Pressure Clusters Based on Image Analysis

For each pressure cluster in the camera image, GravitySpace analyzes the probability of being one of the contact types shown in Figure 12. GravitySpace distinguishes hands, knees, buttocks, and shoes, thereby further distinguishing between heel, tip, and edge of a shoe. These probability dis-

tributions are an essential part of the subsequent pose recognition. Areas covered by furniture pieces are ignored for this classification in order to minimize noise.

In order to classify each pressure cluster, GravitySpace extracts 16 fast-to-compute image features from the respective area in the image, including image moments, structure descriptors using differences of Gaussians, as well as the extents, area, and aspect ratio of the bounding box around the cluster. We trained a feedforward neural network, which assigns probabilities for each type of contact to each cluster.
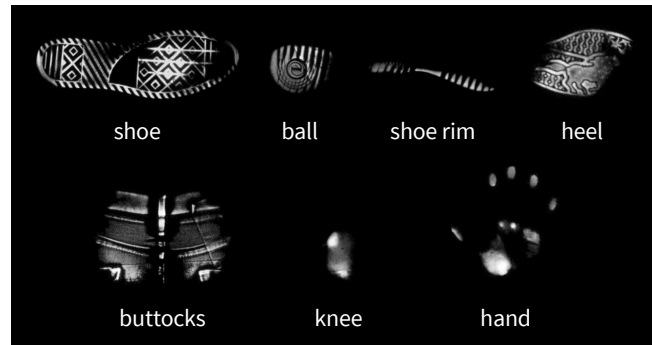

Figure 12: GravitySpace assigns each pressure cluster the probability of being one of the contact types shown above.

Step 4: Identifying Users Based on Shoeprints

Whenever the users' feet are in contact with the ground—for example when standing or sitting, but not when lying—GravitySpace will recognize users by matching their shoeprints against a database of shoes associated with user identities. As users register with both of their shoes, our approach also distinguishes left and right feet.

Due to the large floor area, previous approaches to user identification on multi-touch floors are not sufficient on GravitySpace, such as template matching in *Multitoe* [1], which observed an area of only a 20th of the size.
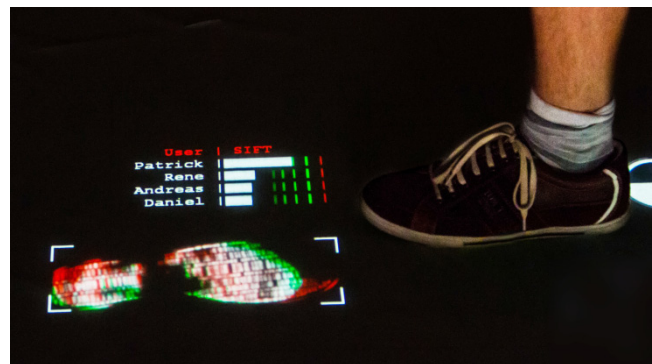

Figure 13: GravitySpace uses SIFT to match detected shoeprints against a database of registered users.

To match shoeprints with the same resolution as previous systems (1 mm per pixel), GravitySpace uses an implementation of SIFT [19] that runs on the GPU. Using SIFT as the feature detector and descriptor algorithm allows us to match shoes with rotation invariance. To identify a user by the shoe, GravitySpace counts the number of features that match in each of the shoe images in the database and the observed

shoeprint as shown in Figure 13. A feature thereby matches if the angular distance between the two descriptor vectors is within close range. As the number of detected features varies substantially between different sole patterns, we normalize the distance (i.e., we divide by the maximum number of features in either observed or database image).

Stitching a Shoe Imprint from a Sequence of Frames

When a user walks on the floor, only a small part of their shoe appears in the image at first and then becomes larger as their shoe sole rolls over the floor from heel to toe. Since the camera consecutively captures images of each such partial shoe imprint, GravitySpace merges all partial observations in successive frames into an aggregated imprint, which allows us to capture an almost complete shoe sole. This concept is also commonly used to obtain a more encompassing finger-print by rolling a finger sideways while taking fingerprints.

Recovering Shoe Orientation Based on Phase Correlation

To predict the location of users' next steps when walking, GravitySpace leverages the orientation of the shoes on the floor. GravitySpace determines shoe orientations directly after matching shoeprints by registering front and back of each database shoeprint with the observed shoe on the floor. Our system transforms both shoeprints into spectrum images and applies log polar transforms to then compute the translation vector and rotation angle between the two shoeprints using phase correlation. All shoes in the database thereby have annotated locations of heel and toes, which happens automatically upon registration by analyzing the direction of walking.

Step 5a: Pose Recognition Based on Spatial Configurations of Pressure Clusters

To classify body poses from the observed pressure imprints, GravitySpace performs pose matching based on the location and classified type of observed pressure clusters. For example, GravitySpace observes the spatial configuration of pressure clusters shown in Figure 14, i.e., the imprints of buttocks, two feet, and two hands as a user is sitting on the floor.

To match a pose, GravitySpace uses a set of *detectors*, one for each pose that is registered with the system. Each detector is a set of rules based on contact types and their spatial arrangement. GravitySpace currently distinguishes five poses: standing, kneeling, sitting on the floor, sitting on cube seat or sofa, and lying on a sofa.

GravitySpace feeds all pressure clusters to all detectors. Each detector creates a set of hypotheses. Each hypothesis, in turn, contains a set of imprints that match the pose described by the detector. For example, hypotheses returned by the sitting detector contain buttocks and two feet. Optionally, there may also be two hands if users support themselves while leaning backwards as shown in Figure 14. Each detector returns all possible combinations (or hypotheses) of imprints that match the pose implemented by this detector. Each hypothesis thus explains a subset of all imprints. We compute the probability of a hypothesis by multiplying the classification probabilities of all contained imprints with a pose-specific prior.
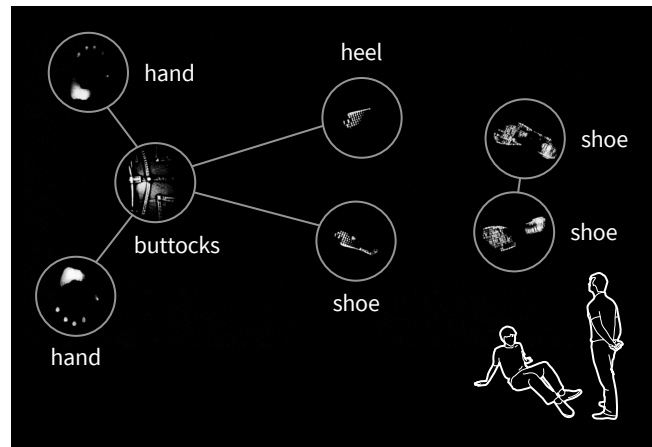

Figure 14: Based on pressure clusters types and their spatial arrangement, GravitySpace recognizes a sitting and a standing user.

From these *individual* hypotheses (explaining a single pose), we compute a set of *complete* hypotheses; each complete hypothesis explains all detected imprints by combining individual hypotheses. We calculate the probability of a complete hypothesis as joint probability of individual hypotheses, assuming that individual poses are independent from each other. We track complete hypotheses over multiple frames using a Hidden Markov Model with complete hypotheses as values of the latent state variable.

Step 5b: Tracking Based on Pressure Distributions

GravitySpace also tracks body parts that are *not* in contact with the floor, such as the locations of feet above the ground while walking or kicking. Our system also tracks general body tilt, for example when a user leans left or right while sitting. This allows for predicting steps before making contact with the floor to reduce tracking latency or to interact with virtual objects on the floor. Obviously, our approach cannot sense events taking place in mid-air, such as raising an arm or changing the gaze direction.

We estimate the location of in-air joints by analyzing the changing centers of gravity within each pressure cluster. We then try to best fit a skeleton to the computed locations of all joints using a CCD implementation of inverse kinematics. GravitySpace finally visualizes the reconstructed body poses with 3D avatars.

Deriving the Location of Feet above the Ground

GravitySpace enables users to interact with virtual objects, such as by kicking a virtual ball as shown in Figure 8. To simulate the physical behavior of virtual objects, GravitySpace first computes the position of feet above the ground. Since a foot is not in contact with the ground when kicking, GravitySpace reconstructs its location by analyzing the changing pressure distribution of the *other* foot, which is on the ground as shown in Figure 15. Our algorithm first calculates the vector from the center of pressure of the cluster aggregated over time to the center of pressure of the current cluster. This vector corresponds to the direction that a person is leaning towards, and is used to directly set the position of the foot in mid-air. We again derive a skeleton using inverse

kinematics, which enables animating the remaining joints of the avatar for output.

Tracking body tilt

To track a user's body tilt, for example when leaning left or right when playing the video game described in the walkthrough, GravitySpace observes multiple pressure clusters as shown in Figure 16. The system first computes the joint center of pressure over all pressure clusters of a user by summing up zero and first order moments of the individual pressure images. We then exploit that the center of pressure directly corresponds to a body's center of gravity projected on the floor. Once the center of gravity is determined, GravitySpace sets the corresponding endpoints of the skeleton's kinematic chains; all other joints then follow automatically based on the inverse kinematic.
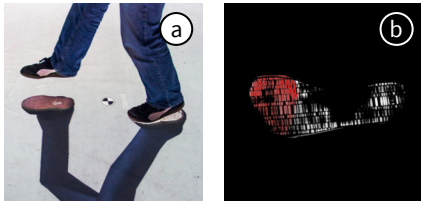

Figure 15: GravitySpace (a) derives the foot location above the floor based on (b) pressure distributions of the other foot.
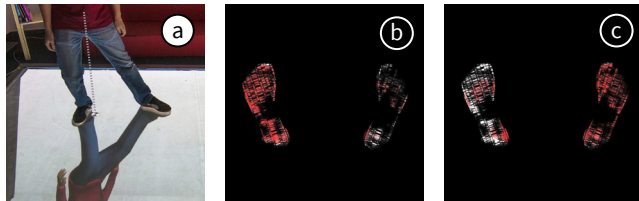

Figure 16: (a) Tracking a user's center of gravity based on (b) and (c) the joint center of pressure of both feet.

EVALUATION

We conducted a technical evaluation of three system components, namely pressure cluster classification, user identification, and pose recognition. In summary, the algorithms of our prototype system allow for (1) distinguishing different body parts on the floor with an accuracy of 92.62% based on image analysis of pressure clusters, (2) recognizing four body poses with an accuracy of 86.12% based on type and spatial relationships between pressure clusters, and (3) identifying 20 users against a 120-user database with an accuracy of 99.82% based on shoeprint matching.

Pressure Cluster Classification

To evaluate pressure cluster classification, we trained a feedforward neural network with data from 12 participants, and tested its classification performance with data from another four participants.

Training Data: We asked 12 participants to walk, stand, kneel, and sit on the floor, in order to collect data of the seven different contact types required for pose recognition, namely hand, shoe (we distinguish between the entire shoe, ball, rim, and heel), knee, and buttocks. In total, we collected 18,600 training samples.

	shoe	ball	shoe rim	heel	buttocks	knee	hand
shoe	1088	1	2	0	0	0	1
ball	17	200	2	18	1	0	12
shoe rim	14	1	297	5	0	0	1
heel	2	83	35	101	0	0	0
buttocks	23	3	0	1	254	3	6
knee	0	17	3	3	0	366	2
hand	7	0	31	21	41	58	459

Figure 17: Confusion matrix for pressure cluster classification.

Test Data: Following the same procedure, we collected data from another four participants, for testing. This resulted in 3,127 samples.

Evaluation Procedure: We manually annotated all training samples to provide ground truth. We then fed the test data into the trained neural network, taking the contact type with the highest probability as outcome. Note that our algorithm does not discard the probability distributions provided by the neural network, but feeds them into the following pose recognition as additional input.

Results: Our approach achieved a classification accuracy of 86.94% for the seven contact types shown in the confusion matrix of Figure 17. If the entire shoe, ball, rim, and heel are grouped and treated as a single contact of type "shoe", as done by the pose recognition, classification accuracy reaches 92.62%.

Pose Recognition

We evaluated our pose recognition implementation with five participants. As pose recognition is based on descriptors of spatial contact layouts, no training data is required for this evaluation.
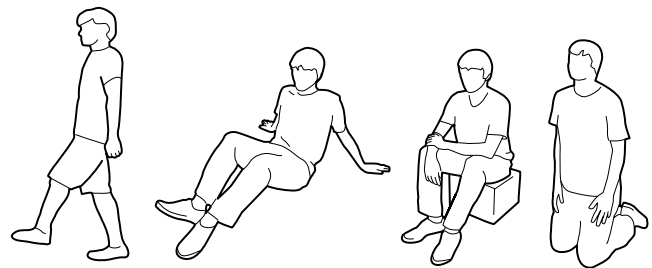

Figure 18: We tested identification of four poses: standing/walking, sitting, sitting on furniture, and kneeling.

Test Data: We collected data from five participants, who each performed the four poses shown in Figure 18, namely standing/walking, sitting on the floor, sitting on furniture, and kneeling. For each participant and pose, we recorded a separate pressure video sequence.

Evaluation Procedure: To provide ground truth, we manually annotated all frames with the currently shown pose. We then ran our algorithm on all frames of the recorded videos, and compared the detected poses to ground truth annotations.

Results: 86.12% (SD=13.5) of poses were correctly identified within a time window of 1.5 s as tolerance. In comparison, FootSee achieves recognition rates of 80% (tested with a single subject) for five standing only activities [44].

User Identification

We determined the user identification accuracy of our implementation with 20 users.

Registration: To populate the user database, each participant walked in circles for about 35 steps on the floor. GravitySpace now selected one left and one right shoeprint for each participant, choosing the shoeprint with the minimum distance in the feature space compared to all other shoeprints of the same participant and foot.

Test Data: After a short break, participants walked a test sequence of about 60 steps. Shoeprints were in contact with the floor for an average of 0.92 s (SD=0.13). Participants then did another round. This time, however, they were instructed to walk as fast as possible, resulting in a sequence of about 70 steps with a lower average duration of 0.38 s (SD=0.11).
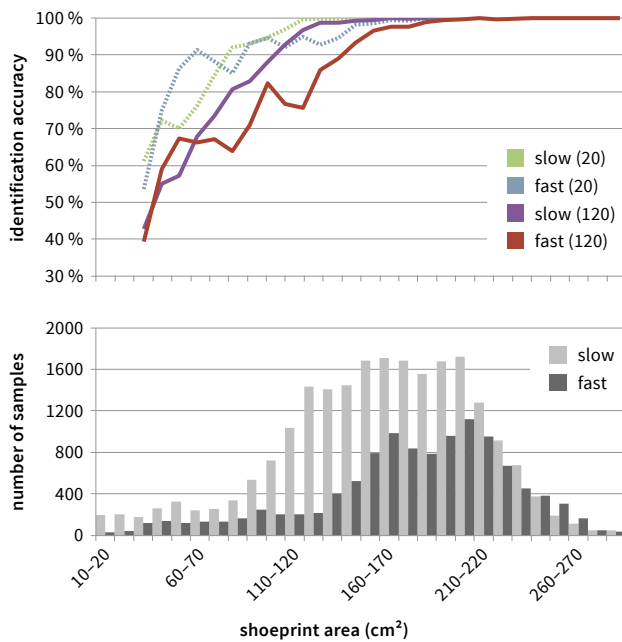

Figure 19: Identification accuracy for walking slow and fast, using databases of 20 and 120 users. Aggregating multiple frames leads to more complete shoeprints and better identification accuracies.

Evaluation Procedure: We evaluated the identification performance by running our algorithms on the recorded test data. Obviously, the slower participants walked, the longer their feet were in contact with the floor, and the more frames were available. The foot part in contact with the floor varied while walking, rolling from heel to toe. As described above, our algorithm reconstructed shoeprints by merging successive pressure imprints. We ran our identification algorithms on all aggregated imprints with an area greater than 30 cm^2, which is the minimum area for discernible shoe contacts as determined during the previous pressure cluster evaluation.

Results: We evaluated the test set against two databases, one containing 20 study participants, and one enlarged with data from 100 additional people (i.e., lab members and visitors). Figure 19 shows the identification accuracy using these two databases for both walking slow and fast. As expected, larger shoeprints aggregated from more frames resulted in better recognition. For the 20-user database, the classification accuracy reached 99.94% for shoeprints with an area between 180 and 190 cm² (the average area of shoeprints). When walking fast, recognition rates slightly dropped to a maximum classification accuracy of 99.19%. This is expected as shoeprints were more blurry.

We then reran the classification against the 120-user database. Our approach correctly identified 99.82% shoeprints when walking slowly, and 97.56% when walking fast. In comparison, Orr et al.'s Smart Floor identifies users based on their footstep force profiles and achieves recognition rates of 93% for 15 users [26]. Qian et al. correctly recognize 94% of 10 users based on gait analysis [31].

Speed: Feature extraction took 47.7 ms (SD = 11.4) per shoeprint, which is independent of the number of registered users. Identification took 251.4 ms (SD = 81.2) using a database of 120 users. Each additionally registered user increases the runtime by 2 ms.

To maintain a frame rate of 25 fps, GravitySpace runs user identification asynchronously. Before identification is completed, users are tracked based on heuristics (e.g., distance and orientation of shoeprints). Once identified, user tracking relies on this information. To reduce delays due to identification, GravitySpace caches recently seen users: new contacts are first compared to this short list of before falling back to the entire participant database.

CONCLUSIONS AND FUTURE WORK

We have demonstrated how to track people and furniture based on a high-resolution 8 m^2 pressure-sensitive floor. While our sensor is limited to sensing *contact* with the surface, we have demonstrated how to conclude a range of objects and events that take place *above* the surface, such as user pose and collisions with virtual objects. We demonstrated how to extend the range of this approach by sensing *through* passive furniture that propagates pressure to the floor. As future work, we plan to combine GravitySpace with other touch-sensitive surfaces into an all touch-sensitive room, and explore the space of explicit interaction across large floors.

ACKNOWLEDGMENTS

REFERENCES

1. Augsten, T., Kaefer K., Meusel R., Fetzer C., Kanitz D., Stoff T., Becker T., Holz C., and Baudisch P. Multitoe: High-precision interaction with back-projected floors based on high-resolution multi-touch input. In *Proc. UIST'10*, 209–218.
2. Baudisch, P., Becker, T., and Rudeck, F. Lumino: Tangible blocks for tabletop computers based on glass fiber bundles. In *Proc. CHI'10*, 1165-1174.
3. Bartindale, T. and Harrison, C. Stacks on the surface: Resolving physical order with masked fiducial markers. In *Proc. ITS'09*, 57–60.
4. Borchers, J., Ringel, M., Tyler, J., and Fox, A. Stanford Interactive Workspaces: A framework for physical and graphical user interface prototyping. *IEEE Wireless Communications, Vol. 9*, No. 6, IEEE Press, December 2002, 64–69.
5. Brumitt, B., Meyers, B., Krumm, J., Kern, A., and Shafer, S. A. EasyLiving: Technologies for intelligent environments. In *Proc. Ubicomp'00*, 12–29.
6. Cruz-Neira, C., Sandin, D.J., and DeFanti, T.A. Surround-screen projection-based virtual reality: The design and implementation of the CAVE. In *Proc. SIGGRAPH'93*, 135–142.
7. Davidson, P., Han, J. Extending 2D object arrangement with pressure-sensitive layering cues. In *Proc. UIST'08*, 87–90.
8. Davies, E. Roy (2005). *Machine Vision: Theory, Algorithms, Practicalities (3rd ed.)*. Amsterdam, Boston: Elsevier. ISBN 9780122060939.
9. Dietz, P. and Leigh, D. DiamondTouch: A multi-user touch technology. In *Proc. UIST'01*, 219–226.
10. F-Scan® System in-shoe plantar pressure analysis http://www.tekscan.com
11. González, R., and Woods, R. *Digital image processing*. Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA ©1992 ISBN:0201508036.
12. Grønbæk, K., Iversen, O. S., Kortbek, K. J., Nielsen, K. R., and Aagaard, L. IGameFloor: A platform for co-located collaborative games. In *Proc. ACE'07*, 64–71.
13. Han, J. Y. Low-cost multi-touch sensing through frustrated total internal reflection. In *Proc. UIST'05*, 115–118.
14. Holz, C. and Baudisch, P. The generalized perceived input point model and how to double touch accuracy by extracting fingerprints. In *Proc. CHI'10*, 581–590.
15. Kientz, J. A., Patel, S. N., Jones, B., Price, E., Mynatt, E. D., and Abowd, G. D. The Georgia Tech aware home. In *CHI'08 EA*, 3675–3680.
16. Krogh, P. G., Ludvigsen, M., Lykke-Olesen, A. Help me pull that cursor. A Collaborative Interactive Floor Enhancing Community Interaction. In *Proc. OZCH'04*, 22–24.
17. LaViola, J. J., Feliz, D. A., Keefe, D. F., and Zeleznik, R. C. Hands-free multi-scale navigation in virtual environments. In *Proc. I3D'01*, 9–15.
18. Leitner, J. and Haller, M. Geckos: Geckos: Combining magnets and pressure images to enable new tangible-object design and interaction. In *Proc. CHI'11*, pp 2985–2994.
19. Lowe, D. Object recognition from local scale-invariant features. *IEEE Comp. Vision 2*. (1999), 1150–1157.
20. Matsushita, N., and Rekimoto, J. *HoloWall*: Designing a finger, hand, body, and object sensitive wall. In *Proc. UIST'97*, 209–210.
21. Molyneaux, D., Izadi, S., Kim, D., Hilliges, O., Hodges, S., Cao, X., Butler, A., and Gellersen, H. Interactive environment-aware handheld projectors for pervasive computing spaces. In *Proc. Pervasive'12*, 197–215.
22. Morris, M.R. Designing tabletop groupware. In *UIST'05 Doctoral Symposium*.
23. Mota, S. and Picard, R.W. Automated posture analysis for detecting learner's interest level. In *Proc. CVPRW'03*, 49–56.
24. Mutlu, B. and Krause, A. and Forlizzi, J. and Guestrin, C. and Hodgins, J. Robust, low-cost, non-intrusive sensing and recognition of seated postures. In *Proc. UIST'07*,–158.
25. Olwal, A. and Wilson, A.D. SurfaceFusion: Unobtrusive tracking of everyday objects in tangible user interfaces. In *Proc. GI'08*, 235–242.
26. Orr, R. J. and Abowd, G. D. The Smart Floor: a mechanism for natural user identification and tracking. In *CHI'00 EA*, 275–276.
27. Paradiso, J., Abler, C., Hsiao, K., Reynolds, M. The Magic Carpet: Physical sensing for immersive environments. In *Proc. CHI'97*, 277–278.
28. Pavlou, M. and Allinson, N.M. Automatic extraction and classification of footwear patterns. In *Proc. IDEAL'06*, 721–728.
29. Pavlou, M. and Allinson, N.M. Footwear Recognition. In *Proc. Encyclopedia of Biometrics'09*, 557-562.
30. Piper, A., O'Brien, Ringel Morris, M., and Winograd, T. SIDES: A cooperative tabletop computer game for social skills development. In *Proc. CSCW'06*, 1–10.
31. Qian, G., Zhang, J., Kidane, A. People identification using gait via floor pressure sensing and analysis. In *Proc. SCC'08*, 83–98.
32. Richardson, B., Leydon, K., Fernstrom, M., and Paradiso, J. A. Z-Tiles: Building blocks for modular, pressure-sensing floorspaces. In *CHI'04 EA*, 1529–1532.
33. Richter, S., Holz, C., and Baudisch, P. Bootstrapper: Recognizing tabletop users by their shoes. In *CHI'12*, 1249–1252.
34. Rosenberg, I. and Perlin, K. The UnMousePad: An interpolating multi-touch force-sensing input pad. In *Proc. SIGGRAPH'09/ACM Trans. Graph.* (Vol. 28, No. 3, Article 65), 1–9.
35. Schmidt, D., Chong, M., and Gellersen, H. HandsDown: Hand-contour-based user identification for interactive surfaces. In *Proc. NordiCHI'10*, 432–441.
36. Srinivasan, P., Birchfield, D., Qian, G. and Kidan, A. A pressure sensing floor for interactive media applications. In *Proc. ACE'05*, 278–281.
37. Streitz, N., Geißler, J., Holmer, T., Konomi, S., Müller-Tomfelde, C., Reischl, W., Rexroth, P., Seitz, P., and Steinmetz, R. i-LAND: An interactive landscape for creativity and innovation. In *Proc. CHI'99*, 120–127.
38. Sugiura, A. and Koseki, Y. A user interface using fingerprint recognition: holding commands and data objects on fingers. In *Proc. UIST'98*, 71–79.
39. Verbunt M, Bartneck C. Sensing senses: Tactile feedback for the prevention of decubitus ulcers. In *Appl Psychophysiol Biofeedback*. 2010 Sep, 35(3):243–50.
40. Villar, N., Gilleade, K., Ramduny-Ellis, D., and Gellerson, H. The VoodooIO gaming kit: A real-time adaptable gaming controller. In *Proc. ACE'06*.
41. Visell, Y., Law, A., and Cooperstock, J. R. Touch is everywhere: Floor surfaces as ambient haptic interfaces. *EEE Trans. Haptics 2, 3 (Jul. 2009)*, 148–159.
42. Weiser, M. The computer for the 21st century. *Scientific American*, Sep. 1991, 66–75.
43. Wilson, A. and Benko, H. Combining multiple depth cameras and projectors for interactions on, above and between surfaces. In *Proc. UIST'10*, 273–282.
44. Yin, K. and Pai, D. FootSee: An interactive animation system. In *Proc. SCA'03*, 329–338