HAVE 2008 - IEEE International Workshop on
Haptic Audio Visual Environments and their Applications
Ottawa, Canada, 18-19 October 2008

A Multi-modal Floor-space for Experiencing Material Deformation Underfoot in Virtual Reality

Alvin W. Law[1], Benjamin V. Peck[2], Yon Visell[1], Paul G. Kry[2], Jeremy R. Cooperstock[1]

[1]Centre for Interdisciplinary Research in Music Media and Technology,
Centre for Intelligent Machines, Dept. of Electrical and Computer Engineering,
E–mail: alvinlaw@cim.mcgill.ca, yon@cim.mcgill.ca, jer@cim.mcgill.ca.

[2]School of Computer Science,
E–mail: benjamin.peck@mail.mcgill.ca, kry@cs.mcgill.ca.

3480 University Street, McGill University,
Montreal, Canada

Abstract – *We present a floor-space design that provides the impression of walking on various terrains by rendering graphical, audio and haptic stimuli synchronously with low latency. Currently, interactive floors tend to focus on visual and auditory feedback but have neglected to explore the role of haptics. Our design recreates these three modalities in a direct manner where the sensors and reactive cues are located in the same area. The goal of this project is the creation of a realistic and dynamic area of ground that allows multiple, untethered users to engage in intuitive interaction via locomotion in a virtual or augmented reality environment.*

Keywords – *interactive floor, deformation models, virtual materials*

I. INTRODUCTION

Virtual reality (VR) stimulates our senses to convince us that the surrounding world is a real, yet mutable, environment. One area that has much room for improvement is that of realistic, interactive locomotion in VR spaces. Presented with a new environment, curiousity naturally encourages people to walk around and explore. Locomotion is a critical part of familiarizing oneself with a new place, not just because it allows us to see around corners, but also because information is gathered through the act of walking itself. The sound of crunching leaves underfoot, or the sensation of gravel shifting as it is walked upon, can inform us about where we are, and aid our decisions during navigation. Our feet are the only parts of our bodies in constant contact with the surrounding environment. This project aims to take advantage of the lifetime of experience we have acquired in this way.

The objective of our work is to generate perceptually realistic simulations of ground areas, by modeling their physically interactive properties. If this effort is successful, we might provide people with the sensation of walking on remote terrains, such as arctic snow or desert sand, as appropriate to specific applications, and to change these sensations dynamically. Our floor design (see Figure 1) builds on prior work in the field, and is especially adapted for multi-modal virtual reality. It is intended to look, sound, and feel realistic, drawing on all three modalities simultaneously, and with sufficiently low latency, to provide an immersive experience.

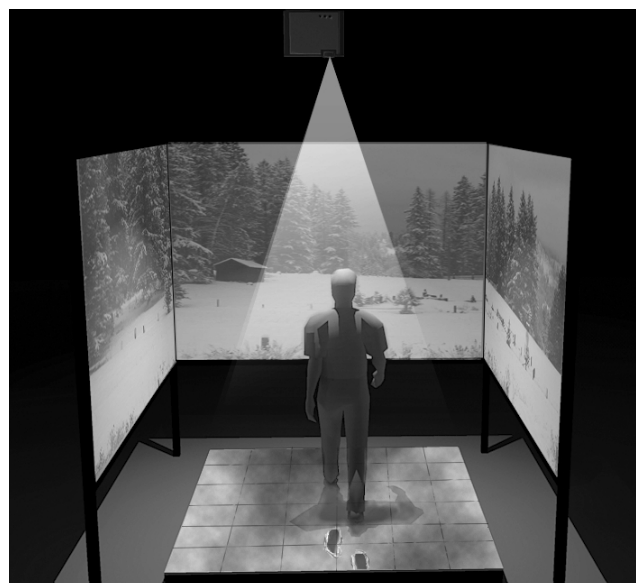

Fig. 1. The floor being constructed for our lab's existing CAVE-like environment adds a fourth dynamic, multi-modal surface.

II. RELATED WORK

As our contribution focuses on integrating and improving both a floor-space and a real-time material-deformation model, relevant work in these respective domains is presented.

A. Interactive Floors

A number of research groups have investigated the development of floors with integrated sensing capabilities, and in some cases, with integrated auditory or visual display capabilities. Many of these use force-sensing resistor (FSR) arrays, arranged in a matrix, to acquire foot-floor contact forces, as in the system we present here. One example is the touch sensitive dance floor by Pinkston [1], which also provides auditory feedback. Other approaches to sensing that have been used in musically interactive floors include the piezoelectric wire network developed by Paradiso et al. [2]. Closer in spirit to the present contribution, Cook's PholieMat [3] uses a sensor mat as a controller for the real-time synthesis of footstep sounds generated by walking on different virtual materials.

Several current systems make use of visual feedback in response to users' footsteps. Commercial products[1] exist that track users with a stereo camera and display advertisements via top-down projection. Using projection and video capture from beneath a glass floor surface, the Wisdom Well [4] creates an educational play area that overcomes the common problem of shadows. The ADA project [5] involves large sensing tiles with integrated color illumination that respond to the patterns of people walking through the space in which it is installed. Although limited in resolution compared to systems utilizing video projection, the project resulted in an interactive floor much larger in scale than previous work.

A comparison of further examples is given in the book by Miranda and Wanderley [6].

B. Deformable Terrain Models

Previous work on simulating deformable terrain can generally be categorized into one of two methods: height-field-based or particle-based approaches. The former represent terrain as a matrix of vertical displacement values, while particle systems simulate the collection of granular elements that make up the terrain. Particle-based models have achieved highly realistic results, but remain impractical for real-time rendering of any significant terrain quantity [7][8]. By simplifying the problem to two dimensions, height fields have been successful in achieving real-time frame rates [9][10]. There are, however, corresponding limitations in simulating certain phenomena accurately, such as the interactions with concave objects. The remainder of this section will discuss literature relevant to our project objectives.

Sumner et al. [9] represent a deformable terrain with a regular grid of cells containing height values. Their simulation consists of three steps: collision detection by ray-tracing, displacement, and erosion. All material in the collision is displaced into the surrounding cells of the height field. The erosion step further spreads material to cells outside of the collision area wherever the slope is sufficiently steep. The type of material being simulated is determined by parameters such as compression coefficient and the terminal angle at which material will slide into a neighbouring cell. For example, sand does not compress but slides at shallow slopes, while snow compresses and only slides at steep slopes. This method yields effective results, but is intended for off-line animation. Our approach is influenced by this work but focuses on interactive real-time deformation.

Onoue et al. [11] introduced the height span map (HS-map), which overcomes the limitations of the standard height field by adding multiple pairs of heights, or *spans*, of material per cell. Ground deformation is calculated by algorithms based on those of Sumner et al. [9]. The resulting model is convincing, but the HS-map must be recalculated whenever the orientation of a rigid object changes; this is computationally expensive and leads to suboptimal display rates (7-14 Hz on a 3 GHz Pentium 4).

Bell et al. [7] use massive collections of particles to simulate discrete elements of ground material. This model does not take into account any cohesive forces, making it effective only for granular materials such as sand. Zhu et al. [8] modify a particle-based water model and incorporate viscosity by tracking the cohesive forces. A unified surface is calculated from the collection of particles, yielding an impressive simulation of completely fluid or semi-fluid bodies. However, neither of these models are currently implemented in real time.

Zeng et al. [10] optimize collision detection with a height field model by rendering the depth information of the rigid object to a texture for comparison against the values in the height field. The velocity with which a rigid object enters the ground is taken into account when applying a deformation. The momentum of material in each cell is tracked, and the slide velocity is calculated using Mohr-Coulomb equations. This creates a much more physically based simulation than previous attempts, as the ground material continues to move dynamically from the momentum transferred by colliding rigid objects. Our method, in contrast, computes a static solution for the terrain at each time step using a simpler approach that runs on the GPU, has higher frame rates, and can handle larger terrain resolutions.

C. Haptic Interaction with Granular Models

Benes et al. [12] present a visual and haptic model of a real-time deformable granular material, based on a height-field model of sand. Interaction with the material is afforded by two different manually operated haptic desktop devices (Force Dimension model Omega and Phantom). In this configuration, the physical device controls a virtual proxy object that mediates the deformation. By comparison, the system presented here is a direct interaction paradigm.

III. SYSTEM DESIGN

The example of snow offers a compelling description of our objectives. When people step in snow, they perceive the *crunch* beneath their feet, they hear the distinctive sound of its compression, and they see the deformation of material around their feet. Our floor is intended to reproduce these three sensations simultaneously and instantaneously.

Our work extends a prototype [13] that successfully displays haptic and audio feedback, using FSRs as inputs. We employ a Vicon motion capture system to track the position and orientation of the walker's feet. This is necessary to provide real-time deformable graphics of the terrain model with sufficient accuracy for convincing graphical feedback. Since

[1] www.displax.com, www.vertigo-systems.com

the imprint is highly dependent on the amount of force being applied, optical tracking alone is insufficient, and FSRs must also be used. Additional benefits arise from the sharing of data between optical and FSR sensing paradigms, as described in Section D.

The current design attempts to address limitations of previous architectures. In particular, we localize all feedback modalities to the floor itself, which we believe is of importance to provide a compelling sense of engagement, all the more so in the case of multiple simultaneous users. The element of haptic feedback is, of course, a major distinguishing aspect of our work. This is motivated in part by past research, which has shown that in the absence of visual or auditory cues, characteristic vibrations of ground materials are sufficient for material identification [14].

A. Haptic-Auditory Sensing and Actuation

To determine position and orientation of the foot, the motion capture system tracks markers attached to the shoe. But actual footsteps are determined far more accurately by the FSRs, which can easily distinguish between a foot hovering just above the floor from one that is in contact. The foot-ground interactions *during* contact are also captured with the floor sensors. This pressure information during a footstep has an important effect on all three feedback modalities.

The floor design consists of thirty six square tiles, each measuring 30.5 cm per side. These are arranged in a 6×6 matrix in the center of our CAVE-like environment.

Our initial prototype used a D-Box motion controller for haptic actuation. Although this allowed for DC motion displacement, latency considerations required us to switch to vibrotactile actuators that vibrate, but do not physically displace the tiles. Discarding the motion controller reduced the sensing-to-actuation latency from 100 ms to less than 20 ms. The reader may wonder whether the familiar experience of walking on complex, compliant ground materials, such as snow, can be simulated by such a device. Fortunately, our past research has shown that in the absence of visual or auditory cues, characteristic vibrations of ground materials are sufficient for material identification [14], even without the benefit of proprioceptive information, for example, the sense of one's feet sinking into snow. As applied weight causes a granular material to compress, a weak reaction force is present compared to the applied force. When the limit of compression is reached, the reaction force equals the applied force, with certain characteristic contact vibrations. Although a vibrating actuator cannot generate the displacement effects during compression, it can reproduce the contact vibrations, which convey a significant amount of perceptual information related to the ground material.The results of that study also suggest that vibrotactile and auditory signals are more important as carriers of information about material identity than is proprioceptive information.

Each tile is attached to a Clark Synthesis Vibrotactile Transducer (Silver Model) with 100 W of vibrating power. These act as the voice coil of a speaker, but fixed to the diaphragm. Low frequencies are perceived as haptic stimuli while higher frequencies are perceived as audio. Although there is a small band of overlap between the two, in general, the modalities can be treated as independent. The tiles are suspended on a square ring of 0.635 cm polyurethane foam to isolate vibrations from the rest of the system, as illustrated in Figure 2. Rubber feet can be installed to isolate the vibrations from the floor. Given the limitations of resolution available from a 6×6 tile array, there will likely be situations in which the walker has both feet on the same tile. In this case, both feet will sense the same vibrations, although these will likely be experienced stronger by the active stepping foot due to the mechanics of pressure distribution between the feet during a natural walking motion.

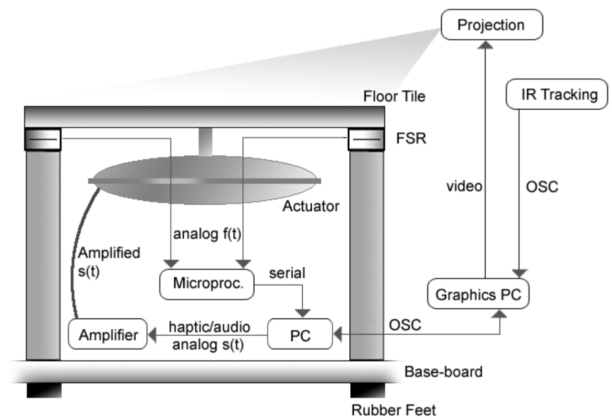

Fig. 2. Side view of a single tile. The components underneath the actuator actually span multiple tiles. Also shown are communication lines between different components. FSR voltage values are denoted by $f(t)$ and the synthesized audio-haptic signal by $s(t)$.

An FSR is recessed in the foam at each corner of the tile, as shown in Figure 3. Taking the relative force readings from the four FSRs allows the location of the center of pressure to be approximated. The sensors can be multiplexed into the analog inputs of an Arduino Diecimila board, supporting up to four tiles of four sensors each, with refresh rates above 250 Hz per sensor. The Arduino is connected by USB to a Mac Mini, which runs the simulation software responsible for haptic and acoustic feedback. USB-based multichannel sound cards are used to transmit dedicated signals to each actuator. Each set of four tile actuators is powered by a four-channel 400 W Class-D audio amplifier. Size concerns influenced the choice of components: we wanted these to fit beneath the actuators, making the height of the floor less than 20 cm, and mostly encapsulated.

B. Haptic-Auditory Synthesis

Haptic and auditory synthesis are accomplished in parallel, as tactile and audio signals are both generated by means of a real-time, lumped signal processing model of interactions with

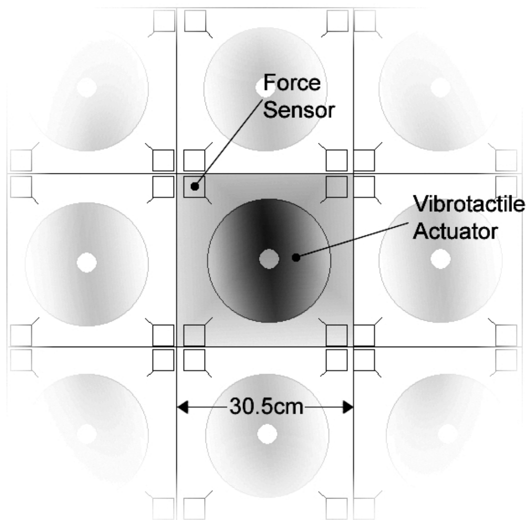

Fig. 3. Top view of the tile array. Each tile contains one actuator in the center and four force sensors located in the corners.

a granular medium, together with a control algorithm matched to the force-sensing floor interface. The model has been described in a prior publication [13], and is based on existing work in footstep sound synthesis by Fontana and Bresin [15]. A measurement scheme has also previously been developed to capture the reaction forces of real world granular ground materials via a custom instrumented shoe. We have used these recordings to identify parameter values of the synthesis model to best match the interactive behavior of these materials.

C. Graphics

Our graphical simulation of the deformable terrain focuses on the interaction between the foot and the ground material. The simulation is displayed on the floor by a ceiling-mounted Hitachi CPX5 projector.

We model the ground with a height field and take advantage of optimizations of Zeng et al. [10] to accelerate collision detection. Our model further optimizes rendering of the height field by using video RAM to store all terrain data and using the GPU to compute lighting normals dynamically. Additionally, we perform all deformations of the height field using only the GPU. Taking advantage of the GPU's stream processing architecture leads to faster simulations and larger terrains than those possible with previous methods. Our deformation algorithm is mapped to the GPU by combining OpenGL shading language programs and render to texture techniques, and is described after our discussion of the model and rendering method, below.

C.1 Model Description and Rendering

The virtual environment consists of a height field and a set of rigid objects, e.g., a model of a foot or shoe, each represented by a 3D triangle mesh. The height field consists of a planar grid of vertices and a floating point texture containing the vertical displacement values for each of the vertices (each pixel of the texture corresponds to one vertex in the plane). We use a texture of dimension 512×512 to simulate a 3.3 m^2 area, which translates to a density of approximately 8 pixels/cm^2. Both the vertex grid and the height map texture are stored in video RAM. This accelerates the rendering process and allows for dynamic computation of lighting normals on the GPU.

The height field uses two height maps. This is required because shader operations are not performed in-place but must be written to a separate texture from that used as input. The first map, or *active* map is used for on-screen rendering and as input for the deformation operation. Terrain deformation is produced in a different sequence of rendering passes (described in Section C.2) which write to the second *inactive* height map. After each simulation iteration, these height maps switch roles (see Figure 4), a process known as "ping-ponging." To render the height field, the ground plane vertices are sent through the graphics pipeline as triangle strips. A vertex shader displaces each vertex in the plane by the amount described in the active height map [16].

Every time the height field is rendered, the vertex shader computes new normals at each vertex by sampling the surrounding height map values. Dynamic computation of normals saves storage space and simplifies deformation operations. Note that the planar field of vertices that makes up the height field mesh remains completely static for the duration of the simulation; only the height maps need to be modified in order to deform the ground.

C.2 Terrain Deformation

The deformation algorithm is triggered once foot pressure is detected on the floor. The first step involves calculating the intersection volume between the foot and the ground. This is done by the collision shader. Following the approach of Zeng et al. [10], the depth buffer information of the foot is rendered to a texture. The depth texture of the foot corresponds to a small subsection of the height map. We only process data in this region, and as such, the computation time for the deformation algorithm is dependent only on the size of the collision area rather than the terrain as a whole. The collision shader compares the depth texture to the active height map. This is done by rendering a single quadrilateral where each pixel corresponds to a sample in the active height map. The output of this shader is rendered to a texture that we refer to as the collision map.

The next step is to calculate the deformation of the ground and produce a new height field. This is done by the deformation shader. Input for this shader consists of the collision map, the active height map, the velocity of the foot, and the compression rate of the material being simulated. This last value defines the percentage of colliding material that is removed from the simulation. The deformation shader initially decreases the height map value by the amount described in the collision map, eliminating the collision with the foot.

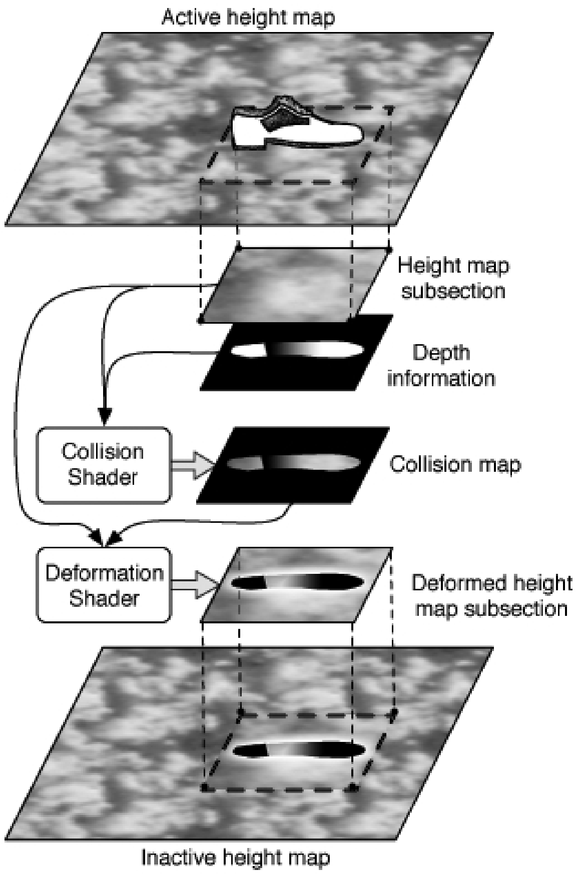

Fig. 4. One simulation step of our graphic model. At the end of each step, the active and inactive height maps are swapped.

To simulate the movement of material out from under the object, a 1D Gaussian blur is applied by the GPU to the collision map data in the direction of movement [17]. The blur width increases with the magnitude of tangential velocity. In addition, a radially symmetric Gaussian blur, independent of the tangential velocity, is also applied, and added to the value sampled from the height map. Depending on the tangential velocity of the object, this blurring either creates a mounding of material next to the foot in the direction of motion (see Figure 5A), or a uniform mounding around the area of collision (see Figure 5B). The resulting deformation map is rendered to the inactive height map in the area defined by the collision, and is displayed when the height maps are swapped. This process is visually effective and provides a reasonable compromise between physical accuracy and computational requirements.

D. Information Integration

Separate software is used to render the graphics and the vibrotactile signals. A Java application reads motion capture data to produce the height field simulation while Max/MSP reads FSR data to generates the acoustic and vibration outputs. These two programs exchange data via OpenSoundControl messages.

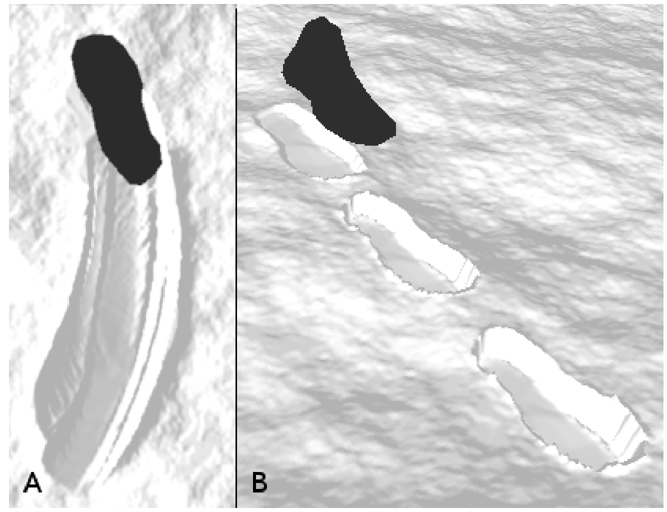

Fig. 5. Screenshots of: A) Deformation due to tangential motion through the terrain. B) Uniform mounding footprints left by the rigid model of a shoe.

While the FSRs provide initial sensor values that determine compression of the ground material, the synthesis model used for vibrotactile and acoustic actuation does not maintain long-term state. It may thus be beneficial to provide information to this synthesis algorithm from the graphical model, as the latter includes a detailed representation of the areas of ground that have already been compressed by footsteps. This would permit modification of auditory and haptic feedback to reflect the level of prior compression to the ground surface. For example, a footstep on compressed snow disperses its energy over a shorter vertical range and time period than on uncompressed snow. In return, the tile system can inform the graphical simulation of applied pressure by the foot to enrich the information provided by the motion capture system, as noted in Section A. This is critically important to achieve realistic effects, as, for example, a light tap of the foot should not produce the same impression as a hard stomp.

IV. ONGOING AND FUTURE WORK

A 2×2 tile prototype of the floor as seen in Figure 6 is presently being adapted and expanded to a 6×6 configuration for deployment in our CAVE-like environment. The effectiveness of the haptic and auditory simulation of snow was already validated through informal studies of the prototype without graphics. Almost all people when asked about what they were walking on comfirmed that it indeed felt like snow.

With regard to performance, we are able to render, on an nVidia GeForce 8800 GTS, a 3.3 m^2 area using a 512×512 height map at 260 Hz. This figure drops to 180 Hz when the

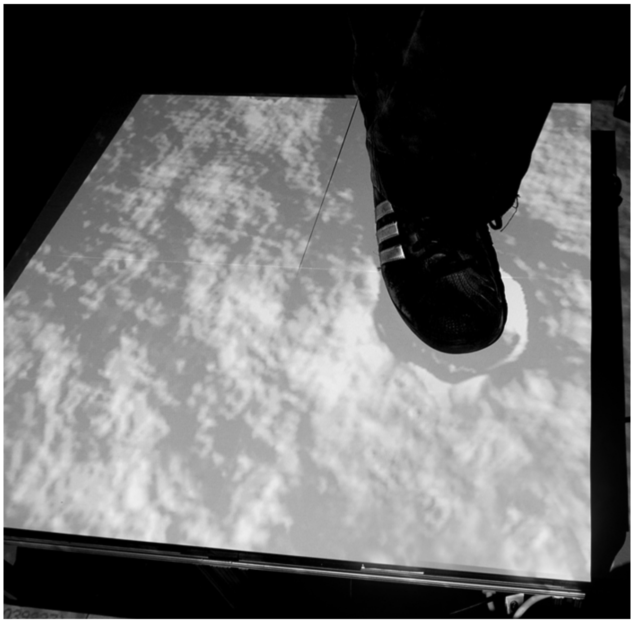

Fig. 6. A footstep causes snow to deform around the shoe (displayed on our prototype).

model is deformed by one foot, and 168-175 Hz with two feet. At 1024×1024 resolution, the height map runs at 90 Hz while rendering and 75 Hz under active deformation.

While the use of simple image filters to recreate deformation-related phenomena is very efficient, some tuning is required to achieve perceptually convincing results. Importantly, additional parameters need to be introduced to the deformation algorithm to support ground materials other than snow. This may also benefit from implementing a more physically based model [10].

We currently project an orthographic image of the terrain onto the floor. Shadows, which may occlude the graphics of interest, are inevitable with overhead projection, and back-projection is infeasible with the sensors and actuators in place. We are therefore considering the use of multiple overlapping projectors, possibly incorporating an active shadow removal process [18] to cope with this problem. The effects of shadows and relative contributions of each modality to users' "sense of immersion" can be assessed using methods such as presence questionnaires, as described by Slater et al. [19].

V. CONCLUSIONS

The system design presented in this paper introduces a novel approach to supporting natural walking activity on virtual terrains. Our prototype applies real-time synthesis to data obtained from both motion capture and force-sensing resistors to provide complementary audio, haptic, and visual stimuli that is highly engaging. While the integration of acoustic and haptic actuation, real-time granular deformation models, and detailed sensing of footsteps, is only now being completed, we are confident that each of these components is sufficiently developed and robust. The resulting fully controllable platform can be used for exploring various questions related to the interpretation of sensory information from the activity of locomotion.

A. Acknowledgments

Funding for this project has been provided by grants from the Natural Sciences and Engineering Research Council (NSERC), Les Fonds Quebecois de recherce sur la nature et les technologies (FQRNT), and the Canada Foundation for Innovation. The support of these agencies is gratefully acknowledged.

VI. REFERENCES

[1] R. Pinkston, "A Touch Sensitive Dance Floor/MIDI Conroller," *Proc. of the ICMC,* 1995.
[2] J. Paradiso, C. Abler, K. Hsiao, M. Reynolds, "The Magic Carpet: Physical Sensing for Immersive Environments," *Proc. of the ACM SIGCHI Conference* March 1997.
[3] P. Cook, "Modelling Bills Gait: Analysis and Parametric Synthesis of Walking Sounds," *Proc. of the Conf. on Virtual, Synthetic and Entertainment Audio,* July 2002.
[4] K. Gronbaek, O. Iversen, K. J. Kortbek, L. Aagaard, "iGameFloor - a Platform for Co-Located Collaborative Games," Proc. of the Intl. Conf. on Advances in Computer Entertainment June 2007.
[5] T. Delbruck, A. M. Whatley, R. Douglas, K. Eng, K. Hepp and P. F. M. J. V. Verschure, "A Tactile Luminous Floor for an Interactive Autonomous Space" *Robotics and Autonomous Systems 55: 433-443.* 2007.
[6] M. M. Wanderley, E. R. Miranda, "New Digital Musical Instruments: Control and Interaction Beyond the Keyboard," *A-R Editions, Middleton, WI,* 2006.
[7] N. Bell, Y. Yu, P. Mucha, "Particle-based Simulation of Granular Materials," *Proc. of the 2005 ACM SIGGRAPH/Eurographics Symposium on Computer Animation* 2005.
[8] Y. Zhu, R. Bridson, "Animating Sand as a Fluid," *ACM SIGGRAPH 2005 Papers* 2005.
[9] R. Sumner, J. O'Brien, J. Hodgins, "Animating Sand, Mud, and Snow," *The Proc. of Graphics Interface '98* 1998.
[10] Y. Zeng, C. Tan, M. Yang, C. Chiang, C. Chang, W. Tai, "A Momentum-Based Deformation System for Granular Material," *Computer Animation and Virtual Worlds* September 2007.
[11] K. Onoue, T. Nishita, "Virtual Sandbox," *Proc. of the 11th Pacific Conference on Computer Graphics and Applications* 2003.
[12] B. Benes, E. Dorjgotov, L. Arns, G. Bertoline, "Granular Material Interactive Manipulation: Touching Sand with Haptic Feedback," *Winter School of Computer Graphics Plzen, Czech republic,* 2006.
[13] Y. Visell, J. Cooperstock, B. Giordano, K. Franinovic, A. Law, S. McAdams, K. Jathal, F. Fontana, "A Vibrotactile Device for Display of Virtual Ground Materials in Walking," *Proc. of the Eurohaptics Conf.* June 2008.
[14] B. L. Giordano, S. McAdams, Y. Visell, J. Cooperstock, H. Yao, V. Hayward, "Non-Visual Identification of Walking Grounds," *Proc. of Acoustics* 2008.
[15] F. Fontana, R. Bresin, "Physics-based Sound Synthesis and Control: Crushing, Walking and Running by Crumpling Sounds," *Proc. of the Colloq. on Musical Infomatics* May 2003.
[16] E. d'Eon, "Deformers," *GPU Gems* 2004.
[17] I. Buck, T. Purcell, "A Toolkit for Computation on GPUs," *GPU Gems* 2004.
[18] S. Audet, J. Cooperstock, "Shadow Removal in Front Projection Environments Using Object Tracking," *Proc. of IEEE Workshop on Projector Camera Systems* 2007.
[19] M. Slater, M. Usoh, A. Steed, "Depth of Presence in Virtual Environments," *Presence: Teleoperators and Virtual Environments,* 3, 1994.