Scrubber: An Interface for Friction-induced Sounds

Georg Essl
Media Lab Europe
Sugar House Lane
Dublin 8, Ireland
georg@mle.media.mit.edu

Sile O'Modhrain
Media Lab Europe
Sugar House Lane
Dublin 8, Ireland
sile@media.mit.edu

ABSTRACT

The Scrubber is a general controller for friction-induced sound. Allowing the user to engage in familiar gestures and feeling actual friction, the synthesized sound gains an evocative nature for the performer and a meaningful relationship between gesture and sound for the audience. It can control a variety of sound synthesis algorithms of which we demonstrate examples based on granular synthesis, wave-table synthesis and physically informed modeling.

1. INTRODUCTION

Friction is a common sound source in everyday life. Squeaking brakes, shuffling noises, brushing teeth, or playing a violin — all these sounds relate in one way or another to friction and hence these sounds can be called friction-induced. The action that causes the friction sound itself also often has a related tactile component, either directly by fingers sliding over a rough surface, or indirectly through a mediating object like a broom or a bow.

In this work we propose one potential implementation of a device that provides control of a broad range of friction sounds while at the same time seeking to retain the tactile and gestural quality of natural actions that induce friction sounds.

The goal of this work is the design of a device for performance of general surface, rough friction interaction which exploits the sensorimotor coupling between action and tactile sensation on the one hand and between action and sonic experience on the other.

Synthesis of friction sounds has a long tradition with respect to the field of computer-based music theory, e.g. analysis and synthesis of the action of a violin bow on a string [22] and on bars [7]. Bowing of conventional strings relies on a rather complicated mechanism of temperature related friction. However, many friction events do not share this characteristic, but are rather more mechanical in nature. They may arise from the roughness of a source, which induces micro-collisions, that are felt and heard as a scrubbing texture and sound. Or the surface may consist of pliable bristles whose mechanical motion with respect to external forces is chiefly responsible for the sensation of the interaction. In either case, granular synthesis remains a natural method for the case of rough and rigid surface friction because little collision impulses and grainy sound make up these kind of sounding events. Granular synthesis has long been an important and widely used compositional technique in Computer Music. Its literature is too extensive to be sufficiently reviewed here — instead we refer the reader to a recent comprehensive exposition by Curtis Roads [18]. For details about granular synthesis and its proposed controllers we refer the reader to the introduction and references of our earlier paper[15]. Friction synthesis and friction based controllers have also been Extensively studied - we refer to Serafin's recent thesis for a comprehensive review [22].

The vast majority of friction controllers relate to the action of the violin bow in some form. Important examples are RBow and BoSSA by Dan Trueman and Perry Cook [23], Diana Young's HyperBow [24] and Charles Nicols' vBow [14].

Recently a body of work on non-bow friction sounds has also emerged. The HyperPuja controller is a controller for the friction of a wooden stick on a Tibetan singing bowl [25]. Rocchesso and co-workers explored non-linear friction models for everyday sounding objects and their control [19]. Rovan and Hayward [20] designed a vibrotactile hand stimulator to simulate the friction related feel of bowing actions.

Pai and coworkers have also explored the relationship between haptic display and sound synthesis in the context of friction [5, 6]. Their AHI is a active force feedback display simulating friction textures and using the input sources to simulate dry friction sounds. Pai *et al* have also explored in some depth the acquisition of dry friction sounds through various means. See [17] and references therein.

In this work, our aim was to design a device that retains tactile qualities of a typical friction event, while allowing flexible interactions and applications for synthesis.

2. DESIGN GOALS

The design goals of the Scrubber are very much inspired by the practical success of an earlier design called PebbleBox [15]. For PebbleBox we were particularly interested in maintaining a loose relationship between the motor action, tactile feel and sonic experience of manipulating physical grains. The basic assumption is that if the relationship between tangible interaction and sonic experience is similar, though not exactly equivalent to real-world experience, the result has the potential to be evocative of a physically valid experience, even though the relationship between audible and tangible properties is loose and approximate.

PebbleBox served as an example of this idea in the realm of granular object manipulation. The Scrubber uses the same paradigm for friction-induced actions and sound.

It has been known for some time that haptic feedback plays an important role in musical performance. Though much of the early evidence of the importance of touch in instrument playing was anecdotal, the advent of computer-

Figure 1: The Scrubber. Note that this work is not sponsored, nor does it endorse this particular eraser brand.

generated haptic or touch feedback has provided a way for researchers to step inside the performer-instrument interaction loop [1, 9, 16, 14]

Since the current goal was to build a controller that couples the feel and sound of friction events, it was important to incorporate into the interface the manipulation of elements that could objectively or subjectively be related to friction. The interaction mode chosen was that of wiping or scrubbing with hand-held objects. These could be for example eraser blocks, sponges or brushes.

For our initial investigation, we chose the eraser block as a place-holder for this general class of object-mediated friction actions. As friction is still physically present in the interaction, the performer should experience a natural tactile sensation of friction coupled to a friction-related sound controlled by the performers actions.

3. CONTROLLER DESIGN

The Scrubber is designed to allow for direct performance of friction-related actions on surfaces in a relatively unconstrained way. The final design can be seen in Figure 1.

The design consists of a gutted white-board eraser (the name of the project is a parody on the name of the particular eraser brand).

The shape was then used to cast a silicone filling with a tubular cavity along its longest extension close to the bottom of the eraser. Into this cavity two actively powered microphone (see Figure 2) were embedded at about one third and three thirds of the distance. The microphone are oriented downward toward the rubbing surface. The purpose of the silicone is to reduce audible artifacts created by grabbing the casing. However, some of these interactions and disturbances are still picked up by the embedded microphone. Additionally, the microphone picks up interactions with the

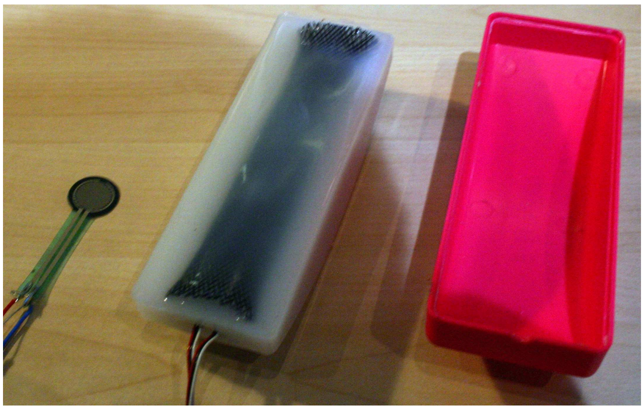

Figure 2: Device components of the Scrubber.

surface. A force sensing resistor is glued to the bottom of the silicone filling to sense overall contact force with the surface.

Typical sounds are the friction-related sound between the device and the surface it is acting upon. Haptic feedback is a result of the direct manipulation of the device conveying the drag on the surface.

4. AUDIO-DRIVEN GRANULAR AND WAVETABLE SYNTHESIS

Live audio based sound manipulation is a known concept. It has for instance been used by Jehan, Machover and coworkers[12, 13], though in their case the relationship between audio event and haptic action was not explicitly retained, as the audio was driven by an ensemble mix of traditional acoustical musical instruments as opposed to the action of any single instrument. As we have discussed in [15] the idea in this work is to extract a parametric version of the sonic qualities of an interaction and use them to resynthesize related sounds. In our earlier paper we discussed this in the context of granular synthesis and offered a simple but functional solution of real-time "granular analysis". Only some friction sounds are well captured by this paradigm. For those where it fits, we have employed the procedure described in [15] and refer the reader there for details. When friction sounds do not have a granular nature, they often relate to rather complicated dynamically sustained sounds. Hence, rather than looking for decay envelopes, as with "granular analysis" of [15], we propose a "friction analysis" which tries to capture features of the sustained sound from the recorded interaction and some supplementary sensor data.

4.1 Frictification Process

To use the raw audio signal as a driver for friction resynthesis, the signal stream needs to be analyzed for friction-like events. The related procedure for granular synthesis was called grainification in [15] and we hence call this process frictification.

The parameters that we considered desirable were event detection in the temporal range of perception ($> .1s$), amplitude envelope measures of a friction event and a measure of spectral content. Additionally we use two channels of audio input to get a measure of direction of the friction action.

The procedure is constrained by the real-time nature of

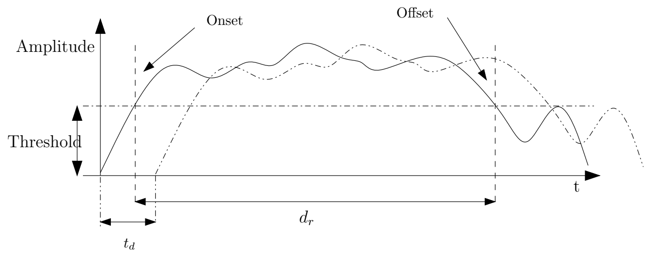

Figure 3: Threshold based grainification scheme. The curve displays an amplitude envelope of an event. d_r is the retrigger delay, preventing detection of new onsets.

the design goal. Firstly, we are bound by causality and hence any consideration for oncoming data translates into delay. Also the amount of processing is bound by the playback buffer length, which in turn translates into delay.

captionThreshold based grainification scheme as used for grain-like friction here. The curve displays an amplitude envelope of an event. d_r is the retrigger delay, preventing detection of new onsets.

The grainification process was kept for friction types which stem from rough and rigid surface interactions. The basic scheme is displayed in Figure 4.1. The frictification process is a modified version of this envelope mechanism for sustained sounds.

Hence, rather than trying to identify onsets on the assumption of rapid decay, we try to detect regions of sustained friction. We employ the following procedure in the current prototype: A very basic onset and offset detection algorithm which also includes a moving short-time zero-crossing average and amplitude envelope. The onsets are detected by thresholding combined with a minimum force reading from the force sensing resistor. The later is used to eliminate false onsets from spurious environmental sounds not related to the friction action. Amplitudes are derived from a moving average after onset initializing with the onset amplitude. This amplitude is assumed to be a sensible measure of friction strength in combination with the data provided by the force sensing resistor. After an onset is detected, the algorithm waits for offsets by observing if the envelope drops below the threshold. If the offset is within one averaging frame, the event is discarded as a non-sustained event and can potentially be directed to a grain synthesizer. The event time d_r is established as the time between onset and offset. Envelope and zero crossing average is set to detect only events that lie in the temporal range of perception ($t > 0.05 - 0.1s$ or alternatively $f < 10 - 20Hz$). For this reason this procedure would not be meaningful for the class of rapidly decaying sounds. This procedure hence constitutes a complement to the grainification process. The relationship of thresholding and event time to a sustained amplitude envelope can be seen in Figure 3.

Finally, some sounds have distinct directional characteristics. The direction of motion is found by the time delay between detected onsets of the two microphone channels. If there is no time delay, the motion is assumed to be normal to the line of the microphones. Otherwise the time delay is taken as a direct measure of directional speed on the axis of the two microphones. The direction and the speed are conveyed as parameters used to allow for directional and speed-dependent playback of friction events.

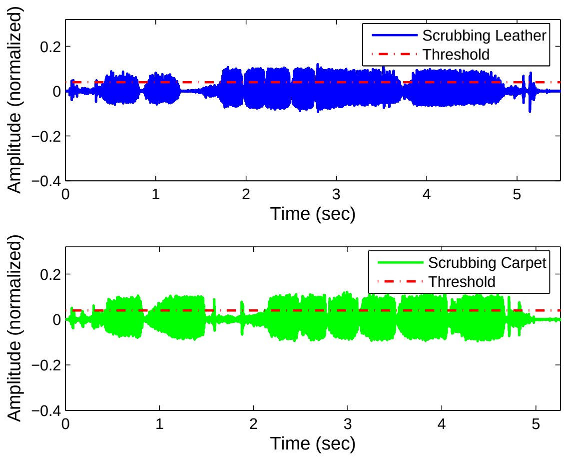

Figure 4: Thresholding of rubbing leather (top) and a carpet (bottom).

This procedure is performed using audio signals from two channels. The inter-channel onset difference is used to estimate direction, as can be seen by comparing the solid and dashed curve in Figure 3.

We found that despite these assumptions and the simplicity of implementation of this procedure reliable sustained event detection and believable control is achieved and hence more advanced methods were not considered at this point.

Figure 4 shows two audio signals as detected, including the thresholds. The first signal rubbing a leather sofa and the second displays a rubbing action on a carpet.

The real-time implementation is based on STK's real-time audio duplexing. We found an input and output buffer size of 128 to work without clicks or missed buffers. This buffer size, at $22050Hz$ corresponds to a basic delay of $11.6ms$. Typically grain estimation windows of 100 samples were used leading to a total delay of around $16.1ms$. Performance measures are taken on a $2.8GHz$ Pentium 4 PC running Windows XP with 512 MB ram and a SigmaTel C-Major Digital Audio device.

5. SYNTHESIS OF FRICTION SOUNDS

We implemented a number of basic synthesis modes. They are all based on basic Wave-table playback and the size, mixing and playback style with relation to the sensory data defines the main difference between them. The plain grain mode is the basic granular synthesis method as described in [15]. This mode is best fitted to rough sources, with the aim of recovering impulsive excitations generated by rough surface interactions. This is also the basic model utilized in the work of Dinesh Pai and co-workers [5].

In order to account for directional property of the interactions, either due to nonlinear properties of the surface or of the virtual actor, we get a notion of velocity and direction from the controller. An example would be bristles as studied for example by Rocchesso and co-worker [19]. Here the direction as sensed by the Scrubber determines the subspace of Wave-table playback. In order to realistically play back complex friction textures, we use continuous variable rate playback. This is equivalent to flexible playback Wave-table

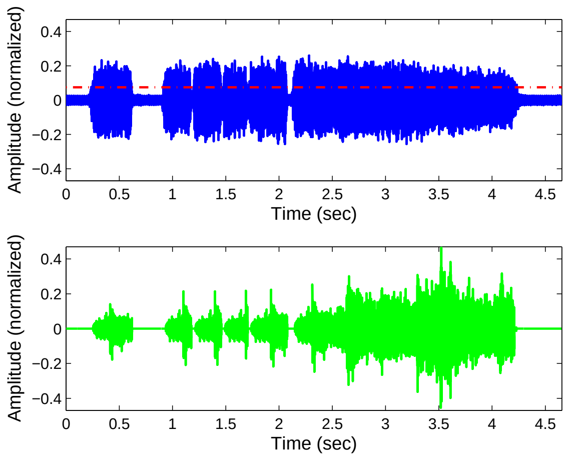

Figure 5: Recorded signal of the Scrubber (top) and frictified response using a sliding garage door sample (bottom) of the complete frictification process.

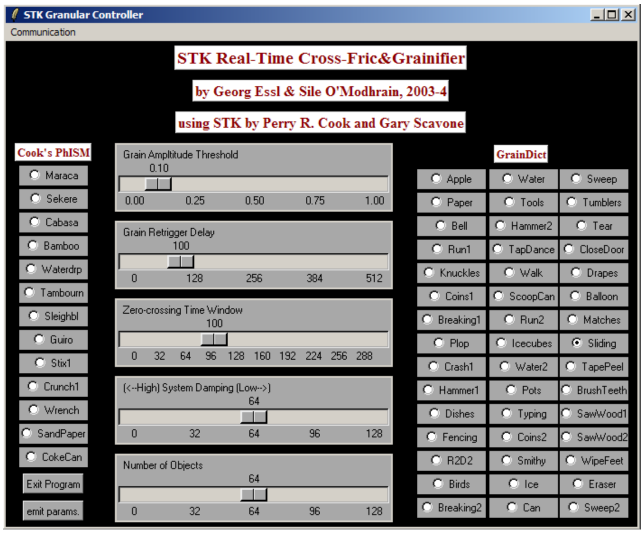

Figure 6: The interface of the Grainification and Synthesis Application GUI implemented in STK.

synthesis.

6. EXAMPLES AND APPLICATIONS

To test the controller in a real application, the extracted data needs to be mapped to sound generation mechanisms. This is the mapping problem, which has seen both theoretical and experimental advances [10; 11; 21, for example].

In principle the sensed data can be mapped arbitrarily. Here we consider the application of our controller design to three types of sound synthesis methods. The first two are based on recorded dictionaries of environmental sounds. The third uses parametric physically informed models developed by Perry Cook [2, 3, 4] and the four are physical models of sustained sounds using waveguides and banded waveguides [8].

6.1 Recorded Environmental Sounds

We implemented a prototype grain and friction sound dictionary based on recordings of natural sounds. 15 friction sounds were added to an existing dictionary of 30 grains. A single dictionary entry could consist of one to 18 recordings of similar but distinct events. More recordings were used when similar interactions led to different sonic experiences, as for example the brushing of teeth has distinct directional features in the sound or where the detail of the interaction is hard to control and hence leads to variation as in the case of tearing, or peeling.

The recordings are played back based on the appropriate friction or granular parameters in the frictification or grainification process. The onset time triggers a variable playback event with the playback amplitude defined by the detected onset and amplitude. Sometimes the playback rate, as a measure of the events overall frequency, was varied with the average zero crossing during a valid detected event. The relationship between recorded friction sounds and final sound using recordings of sweeping wood using the Scrubber can be seen in Figure 5.

6.2 Physically Informed Parametric Models

In order to explore parametric models, we used Perry Cook's shaker based granular synthesis as implemented in his STK software [4] (see the left button row in Figure 6).

Here the mapping of grain onset time and amplitude relates to time and amount of energy infused into the physically inspired model. The zero-crossing average is mapped to the center resonance frequency of the models. These models have inherent stochastic variability. Also some do respond more immediately to energy infusion than others. This does affect the perception of playability, and in general a strong correlation of energy infusion to granular events is desirable. For details on the parametric model synthesis we refer the reader to [2, 3, 4].

7. CONCLUSION

We have here proposed the Scrubber as a new interface for friction induced sound. It is a simple, low cost design with a wide range of friction related gestures. By means of the nature of its parameters it suggests natural mappings to sound synthesis algorithms. The main advantage is the maintenance of familiar relationships of gesture, tactile perception and sonic response for the performer and the audience. Moreover, as with the PebbleBox, the system allows expressive gestural nuances to be passed through to the control of the resulting audio.

The current design is an early prototype. An number of additions would be desirable to address current shortcomings. By choice of microphones as sensors, undesirable noises are potential sources for misrecognitions. This has already been addressed in part by the addition of a force sensing resistor, but additional sensors for capturing specific components of gestures would be desirable. For example, the detection of motion direction is currently limited, as the component of only one axis in the plane can be sensed. We had planned to include accelerometer channels — a plan that had to be postponed due to the rapid redeployment of the workforce.

Finally, the current design is tethered. A wireless design would increase mobility and portability of the controller and thereby support a much richer range of performance opportunities.

This is not the first proposed friction controller. We feel, however, that we propose here a particularly simple and easy design with a broad range of applications. Additionally the design inherently keeps aspects of the interaction intact that are meaningful for natural object-mediated friction sounds and hence by design maintains a familiarity that may be lost in abstracted controllers or active feedback designs.

Acknowledgments. We would like to thank Lily Shirvanee for preparing the silicone mold for us. Stephen Hughes was tremendously helpful with electronics questions and provided his analog-to-MIDI box for prototyping of the FSR input. Mike Bennett and Andrea Chew gave critical input during discussion. This work would not have been possible without the structural support of Media Lab Europe, which will close operations the day after this paper is submitted.

8. REFERENCES

[1] C. Cadoz, L. Lisowski, and F. J.-L. A Modular Feedback Keyboard. In *Proceedings of the International Computer Music Conference*, Glasgow, 1990.

[2] P. R. Cook. Physically Informed Sonic Modeling (PhISM): Synthesis of Percussive Sounds. *Computer Music Journal*, 21(3):38–49, 1997.

[3] P. R. Cook. Toward Physically-Informed Parametric Synthesis of Sound Effects. In *Proceedings of the 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA-99)*, pages 1–5, New Paltz, NY, October 17-20 1999.

[4] P. R. Cook. *Real Sound Synthesis for Interactive Applications.* A K Peters, Ltd., July 2002.

[5] D. K. DiFilippo, D. Pai. Contact Interaction with Integrated Audio and Haptics. In *Proceedings of the International Conference on Auditory Display (ICAD)*, 2000.

[6] D. K. DiFilippo, D. Pai. The AHI: An Audio And Haptic Interface For Contact Interactions. In *Proceedings of ACM UIST (13th Annual ACM Symposium on User Interface Software and Technology)*, San Diego, CA, November 5-8 2000.

[7] G. Essl and P. R. Cook. Measurements and efficient simulations of bowed bars. *Journal of the Acoustical Society of America*, 108(1):379–388, 2000.

[8] G. Essl, S. Serafin, P. R. Cook, and J. O. Smith. Theory of Banded Waveguides. *Computer Music Journal*, 28(1):37–50, 2004.

[9] R. B. Gillespie. *Haptic Displays of Systems with Changing Kinematic Constraints: The Virtual Piano Action.* PhD thesis, Stanford University, 1996.

[10] A. Hunt, M. M. Wanderley, and M. Paradis. The importance of parameter mapping in electronic instrument design. In *Proceedings of the 2002 Conference on New Instruments for Musical Expression (NIME-02)*, pages 149–154, Dublin, Ireland, May 24-26 2002.

[11] A. Hunt, M. M. Wanderley, and K. R. Towards a Model for Instrumental Mapping in Expert Musical Interaction. In *Proceedings of the International Computer Music Conference (ICMC-00)*, pages 209–212, Berlin, Germany, August 27-September 1 2000.

[12] T. Jehan, T. Machover, and M. Fabio. Sparkler: An audio-driven interactive live computer performance for symphony orchestra. In *Proceedings of the International Computer Music Conference*, Göteborg, Sweden, September 16-21 2002.

[13] T. Jehan and B. Schoner. An audio-driven, spectral analysis-based, perceptual synthesis engine. In *Proceedings of the 110th Convention of the Audio Engineering Society*, Amsterdam, Netherlands, 2001. Audio Engineering Society.

[14] C. Nicols. The vBow: Development of a Virtual Violin Bow Haptic Human-Computer Interface. In *Proceedings of the 2002 Conference on New Interfaces for Musical Expression (NIME-02)*, pages 29–32, Dublin, Ireland, May 24-26 2002.

[15] S. O'Modhrain and G. Essl. PebbleBox and CrumbleBag: Tactile Interfaces for Granular Synthesis. In *Proceedings of the International Conference for New Interfaces for Musical Expression (NIME)*, Hamamatsu, Japan, 2004.

[16] S. O'Modhrain, S. Serafin, C. Chafe, and J. O. Smith. Qualitative and Quantitative assessments of the Playability of a Virtual Bowed String Instrument. In *Proceedings of the International Computer Music Conference*, Berlin, 2000.

[17] D. K. Pai and P. R. Rizun. The WHaT: a Wireless Haptic Texture sensor. In *Proceedings of the Eleventh Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, March 22-23, 2003.

[18] C. Roads. *Microsound.* MIT Press, Cambridge, Massachusetts, 2001.

[19] D. Rocchesso, F. Avanzini, M. Rath, R. Bresin, and S. Serafin. Contact Sounds for Continuous Feedback. In *Proceedings of the International Workshop on Interactive Sonification*, Bielefeld, Germany, January 2004.

[20] J. Rovan and V. Hayward. Typology of Tactile Sounds and their Synthesis in Gesture-Driven Computer Music Performance. In M. M. Wanderley and M. Battier, editors, *Trends in Gestural Control of Music*, pages 389–405. IRCAM, Paris, France, 2000.

[21] J. B. Rovan, M. M. Wanderley, S. Dubnov, and P. Depalle. Instrumental Gestural Mapping Strategies as Expressivity Determinants in Computer Music Performance. In *Proceedings of Kansei - The Technology of Emotion Workshop*, Genova, Italy, October 3-4 1997. Available online at `http://www.ircam.fr/equipes/analyse-synthese/wanderle/Gestes/Externe/kansei_final.pdf`.

[22] S. Serafin. *The sound of friction: real-time models, playability and musical applications.* PhD thesis, Stanford University, June 2004.

[23] D. Trueman and P. R. Cook. BoSSA: The Deconstructed Violin Reconstructed. In *Proceedings of the International Computer Music Conference (ICMC)*, pages 232–239, Beijing, China, October 22-27 1999.

[24] D. Young. The Hyperbow Controller: Real-Time

Dynamics Measurement of Violin Performance. In *Proceedings of the 2002 Conference on New Interfaces for Musical Expression (NIME-02)*, pages 65–70, Dublin, Ireland, May 24-26 2002.

[25] D. Young and G. Essl. HyperPuja: A Tibetan Singing Bowl Controller. In *Proceedings of the 2003 International Conference on New Interfaces for Musical Expression*, pages 9–14, Montreal, Canada, May 22-24 2003. McGill University.