From Driving to Expressive Music Performance: Ensuring Tempo Smoothness

Jie Liu
Epstein Dep of Industrial & Sys Eng
Integrated Media System Center
University of Southern California
Los Angeles, California 90089, USA
jieliu@usc.edu

Elaine Chew
Epstein Dep of Industrial & Sys Eng
Integrated Media System Center
University of Southern California
Los Angeles, California 90089, USA
echew@usc.edu

Alexandre R. J. François
Computer Science Department
Viterbi School of Engineering
University of Southern California
Los Angeles, California 90089, USA
afrancoi@usc.edu

ABSTRACT

This paper focuses on the mapping strategies in the interface design for the Expression Synthesis Project (ESP). The goal of ESP is to use the metaphor of driving to allow non-experts to interactively create expressive renderings of music pieces. In ESP, the road represents the music, with bends and straight segments representing places where one might slow down and speed up respectively. The user controls the tempo (rate of beats) and dynamics (amplitude) by the way s/he drives through the road. Our design objective in the ESP environment is to define mapping strategies that guarantee tempo smoothness, a hallmark of expert and practiced performances, under all driving conditions. Because the road has non-zero width, several strategies exist for mapping the car's position on the road to the musical score. We propose mapping strategies that ensure tempo smoothness, and provide mathematical and empirical proofs for the success of these strategies.

Categories and Subject Descriptors

H.5 [Information Interfaces and Presentation]: User Interfaces; H.5 [Information Interfaces and Presentation]: Sound and Music Computing

General Terms

Performance, Design, Human Factors

Keywords

Music Performance Rendering, Musical Expression, Driving Interface, Virtual Environment, Tempo Smoothness, Minimum Jerk Model

1. INTRODUCTION

The Expression Synthesis Project (ESP) [[3]][[5]] aims to create a system that takes the metaphor of driving (controlling a car on a road), and uses it to enable non-experts to create their own expressive renderings of a piece of music, without having to first master a musical instrument. Almost everyone will find that driving a car is relatively easy, but the same cannot be said for playing an instrument. The interface not only gives non-experts access to high-level expressive decisions, it also allows expert musicians (including composers) to experiment with expressive choice without having to first master the notes in a piece of music, or the instrument.

Tempo smoothness is an established requirement of expert performances [[2]]. Lack of control in forming smooth tempo trajectories that lead to discontinuities in the tempo graph are telltale signs of an inexperienced performer. It is common knowledge that humans prefer smooth transitions in the handling of an automotive. By using a driving interface on ESP, we capitalize on the natural analogies between automotive control and smooth musical timing.

In ESP, the tempo of the rendered performance is linked to the velocity of the car. One of the main design goals is to ensure that all ways of driving lead to interpretations with smooth tempo change. This paper presents our design strategies for mapping car motion to musical tempo in such a way as to ensure tempo smoothness in ESP. We provide analytical proofs of tempo smoothness, and demonstrate empirically that the strategies lead to smooth tempo output.

We compare our proposed virtual radius mapping strategies (under interior and boundary conditions) with the rudimentary centerline mapping strategy. We also subject the system to extreme conditions of erratic driving to test the robustness of the mapping strategies. We show quantitatively that tempo smoothness is ensured under all conditions when we employ the virtual radius mapping strategy. This smoothness guarantee allows users to freely explore interpretations of their hearings of the piece through the interface.

2. RELATED WORK

Musical expression is the result of the manipulation of tempo, loudness and articulation. The history of creating expressive performances from existing musical scores using machinery can be traced back to the invention of the player piano in the late 1800's [[4]]. Until now, inventors continue to experiment with instruments or devices for creating expressive performances. Their inventions include: gesture-controlled conducting systems [[13]][[15]], and tapping-controlled systems [[11]][[14]].

This paper focuses on tempo smoothness. Mathematically, a smooth function is one that has continuous derivatives up to

Conference'04, Month 1–2, 2004, City, State, Country.

some desired order over a domain range. Smoothness is important in many applications such as robotics and signal processing. The evaluation of smoothness is not obvious when one considers practical applications. A trajectory is considered smooth if it does not change suddenly in time. Hu et al [[10]] introduced a method to analyze the smoothness of speech signals by approximating the trajectory using a lower-order polynomial function. Harris [[8]] used a Fourier transform method to analyze human movement trajectories. The Minimal Jerk Model is a classic but popular method for smoothness analysis [[12]][[18]]. In this paper, we use the jerk value to evaluate the smoothness of the tempo curve in our empirical results.

The remainder of this paper is organized as follows: we provide information on the background of ESP, its system graph and implementation. Then, we describe the road design in ESP, and introduce the mapping strategies for ensuring smooth tempo change. Following the conceptual presentation of the design strategies, we present mathematical proofs of tempo smoothness. Next, we demonstrate that the strategies work in practice by showing some empirical results. Finally, we present the conclusions of this research.

3. BACKGROUND AND IMPLEMENTATION OF ESP

The ESP interface consists of a driving wheel and pedals, and a computer system, including a screen that shows the road and the dynamically changing views from a vehicle in motion, and speakers for rendering the sounds. Figure 1 shows the MOMO driving wheel and pedals that we use in this implementation of the ESP system, and the view of the road. The pedals act as a force loading system for controlling the speed of the car. The road is created from the score, and the user, who is the driver, controls the way in which the road is traversed. The speed of the car corresponds to the tempo of the music, and loudness is correlated with the acceleration as the piece unfolds over time. The user senses the car's dynamics by seeing the moving visuals on the screen, and by listening to the real-time rendering of the music. The sensory feedback helps the user make his/her decision at the next moment in time.

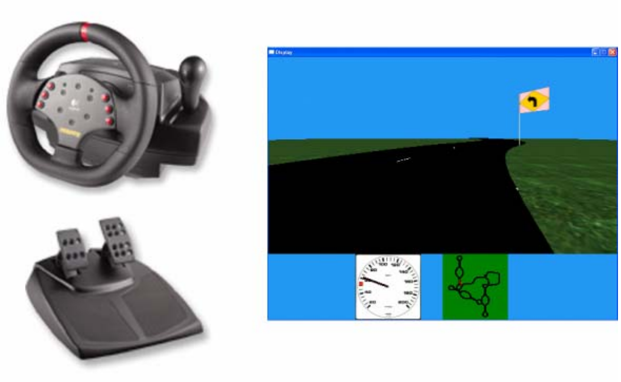

Figure 1 Driving wheel, paddles and visual interface

The ESP system architecture is specified using the Software Architecture for Immersipresence (SAI) framework [[6]] (Francois, 2004), and implemented in C++ using the Modular Flow Scheduling Middleware (MFSM) [[7]], an open source architectural middleware that provides code support for SAI's architectural abstractions. MFSM also offers a number of useful open source functional modules. Using some of these modules significantly reduced the coding effort and allowed us to focus on ESP-specific aspects during the development of the ESP prototype.

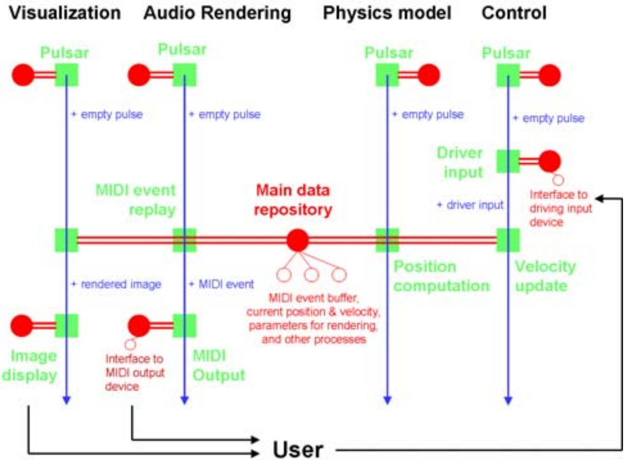

Figure 2 MFSM graph of ESP

Figure 2 shows the conceptual level specification of ESP, expressed in the SAI notation. SAI distinguishes between persistent and volatile data. Persistent data is held in sources (red disks), while volatile data flows down streams (thin blue arrows) in the form of pulses. Pulses travel down streams through processing cells (green squares). The central source holds persistent data structures that are accessed independently by cells placed on independent streams. These data structures include a 3D model of the virtual environment, a dynamic model of the car, and a MIDI representation of the input music piece (list of MIDI events). Each stream defines a functional subsystem, handling (from left to right on the figure) visual rendering, aural rendering, dynamic evolution of the car model, and user control of the car via the input device.

The car's dynamics in ESP are inspired by the principles outlined in [[1]]. The car's acceleration at time t can be described as three components: the lateral, $a_x(t)$, longitudinal, $a_y(t)$, and angular $a_\omega(t)$, accelerations, which are linked to current force input from the pedals, $F(t)$, and the angle of the wheel, $\delta(t)$, and the previous lateral and longitudinal velocities, $v_x(t–1)$ and $v_y(t–1)$ respectively, and the previous frictional force on the front and rear tires, $F_{xf}(t–1)$ and $F_{xr}(t–1)$ respectively in the following way:

$$a_x(t) = F_{xf}(t-1)\cos\delta(t) + F_{xr}(t-1) \\ - \operatorname{sgn} v_x(t-1) c_D v_x^2(t-1) - c_R v_x(t-1)\,, \quad (1)$$

$$a_y(t) = \frac{F(t)}{m} - \operatorname{sgn} v_y(t-1) c_D v_y^2(t-1) - c_R v_y(t-1)\,, \quad (2)$$

$$a_{\omega}(t)=\frac{l_f F_{xf}(t-1)-l_r F_{xr}(t-1)}{I} \tag{3}$$

where c_f is the cornering stiffness of the front tires, c_r is corning stiffness of the rear tires, c_D is the factor for air resistance, and c_R is the factor for rolling resistance. m, I, l_f, l_r *and* h are constants. In this implementation of ESP, we use the values: c_f=–5.20/*rad*, c_r=-5.0/*rad*, c_D=1.5m^{-1}, c_R=5.0s^{-1}, m=1500*kg*, I=1500$kg*m^2$, l_f=1.0*m*, l_r=1.0*m* and h=1.0*m*.

We restrict the control range for the wheel and pedal force as: -0.2*rad* $\le \delta \le$ 0.2*rad* and -12000*N* $\le F \le$ 6000*N*.

All coordinates in the above system are given with respect to the car body. In order to get the absolute coordinates of the car, [$X(t)$, $Y(t)$, $\theta(t)$], we perform the following transformation to get the absolute velocities from the relative velocities:

$$\begin{bmatrix} \dot{X}(t) \\ \dot{Y}(t) \\ \dot{\theta}(t) \end{bmatrix} = \begin{bmatrix} \cos\theta(t) & \sin\theta(t) & 0 \\ -\sin\theta(t) & \cos\theta(t) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_x(t) \\ v_y(t) \\ \omega(t) \end{bmatrix}, \tag{4}$$

where ω is the angular velocity, which can be calculated directly from the angular acceleration, $a_{\omega}(t)$.

As can be seen, at each sampling time we calculate the dynamics of the car, and get a value for the corresponded instantaneous tempo. Our goal is to make the instantaneous tempo vary smoothly.

4. ROAD DESIGN

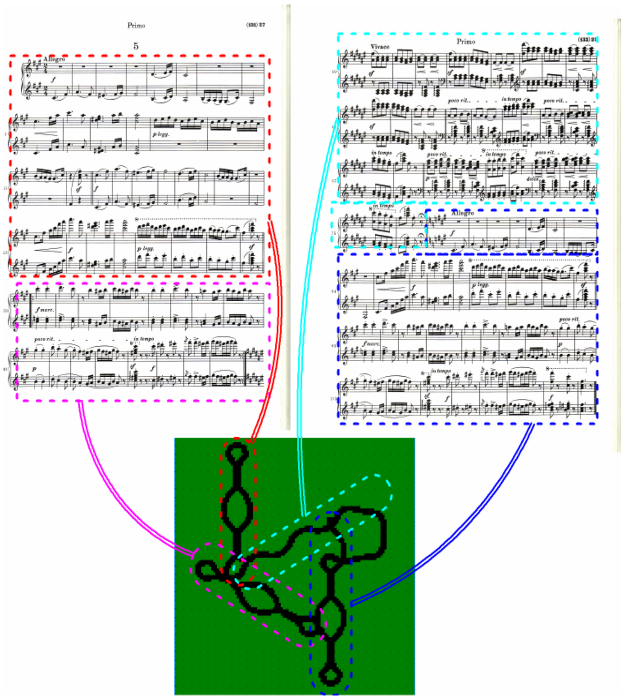

Figure 3 Road Design

In ESP, the road represents an obstacle course that guides the user to make reasonable expressive choices for the piece of the music. Bends in the road suggest to the user that s/he might think about slowing down at these points. A sharp curve indicates a more significant slowdown, while a slight dip in tempo is suggested by a gentler curve. Straight sections suggest to the user that one can speed straight ahead.

Our main test piece in this paper is Brahms' *Hungarian Rhapsody No. 5*. Because performers tend to slow down at phrase endings, the road contains bends at these junctures in the music. Phrase endings of larger sections map to larger bends, and minor phrase endings map to small bends. Tempo slowdowns indicated by "poco rit. …" in the score are mapped to extremely sharp bends, and returns to the regular tempo indicated by "in tempo" in the score are mapped to straight sections. We restrict the turns to alternate left and right to create interest. By following these principles, we create the road for Brahms' *Hungarian Dance No.5* as shown in Figure 3

The AABA structure of the piece, outlined by red, pink, green, and blue dotted lines respectively, is revealed in the repeated structures of the road. The phrase symmetries are reflected in the symmetric patterns within the A sections. At the extreme end of each A section road, we have a very sharp curve that turns for more than 180 degrees, showing physically that the melody is returning.

The bends in the road make it easy to comprehend the structures in the music, but on the other hand, they also make driving near the center of the road difficult. But a linear design could encourage some drivers to simply accelerate for the entire duration of the piece, and thus create a musically uninteresting interpretation.

The road is created from pieces of a varying curvatures and lengths. This piecewise construction allows for quick and easy generation of new roads. It also allows for future expansions of the application to allow the user to piece together their own roads for creating their hearings of the piece.

In ESP, we restrict the car to stay inside the road. From experience, if the car is allowed to stray away from the road, there can be times when it is traveling backwards with respect to the road. Since we are not exploring the performance of music backwards, we do not give the driver the option to go off-road and possibly backwards.

In the above conditions, using a direct mapping of car speed to music tempo, for example, by projecting the car position to the centerline, results in tempo discontinuities. These discontinuities occur at the boundaries between pieces of the road, and when the car hits a side of the road. The next two sections will detail our mapping strategies for creating smooth tempo transitions under the stated conditions

4.1 Mapping Strategy for Boundary Condition 1: Joints

One might ask, why not map the actual speed of the car directly to the tempo of the music? This would work only if the road has a width of 0. For the same road segment, different paths within that segment will be of different lengths, whereas that segment corresponds to a constant segment of music. Then, the most obvious strategy is to project the position of the car to a fixed line, such as the center of the road. Distance along this line represents beats passed in the music score. This centerline projection (CLP) strategy is shown diagrammatically in Figure 4(a).

In Figure 4 (a), the arc *AB* represents the centerline of a curved section of the road. Suppose the car is at position *E* at time *t*, then at position *F* at time $t + \Delta t$. According to the projection strategy, E is mapped to the point *C*, and *F* to *D*. Hence, the arc *AC* represents the distance traveled within the segment when the car reaches point *E*, and *AD* is the distance traveled when the car reaches point F.

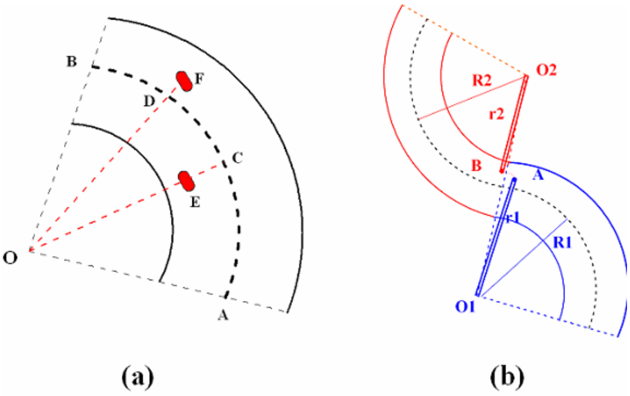

Figure 4 Centerline projection strategy: (a) concept, (b) tempo discontinuity at section boundary

There exists a problem with this basic CLP mapping strategy. Even when the car is moving smoothly, a discontinuity occurs in the tempo when it crosses from one piece of the road to the next. An example is shown in Figure 4(b). Suppose the car is driving from point *A* to point *B*, cross two segments with centers and radii of curvature, O_1 and O_2, and R_1 and R_2 respectively. Let the distance of *A* and *B* to their respective segments' origins be r_1 and r_2 respectively. Suppose the velocity stays constant between *A* and *B*, and is parallel to the corresponding tangents to the road at each time. The projection mapping results in a music tempo $v{*}R_1/r_1$ at *A*, and $v{*}R_2/r_2$ at *B*. From the picture, we can see that $R_1/r_1 < 1$, but $R_2/r_2 > 1$, thus resulting in a sudden change in the multiplicative factor when the car crosses the section boundary even when the car's velocity remains constant. The only exception to the case occurs when the car traces exactly the centerline.

To remove this discontinuity, we introduce the concept of a virtual radius that forces the mapping ratio at either ends of a segment to be forced to 1. At the center of the segment, the virtual radius has the same value as the actual radius; the virtual radius goes to infinity at either ends of the segment. Suppose the piece of road spans an angle α, and that α' has been traversed so far, the formula we use to calculate the present radius of curvature is:

$$R(\alpha') = R\left[1 + \tan\left(\frac{\pi|\alpha' - 0.5\alpha|}{\alpha}\right)\right], \qquad (5)$$

We choose the center of origin to be on the line through *O* and *S* for the first half of the curve, and *O* and *E* for the second half of the curve. Other choices are also possible.

Consider the example shown in Figure 5(a), where a car is at point *A* at time *t*, and at point *B* at time $t + \Delta t$. Performing centerline projection using the origin *O* would result in *A* and *B* being mapped to *D* and *F* respectively. Using Equation (5) to calculate the instantaneous radii, we can find the virtual origins O_A and O_B, so that *A* and *B* map to *C* and *G* instead. Note that Equation (5) is symmetric around $\alpha/2$, and goes to infinity as α' approaches 0 or α. Hence, at the segment boundaries, even though the car may not be on the centerline, as in Figure 4(b), the ratios R_1/r_1 and R_2/r_2 would both be almost 1, thus guaranteeing that the ratio will hardly change when the car crosses the boundary.

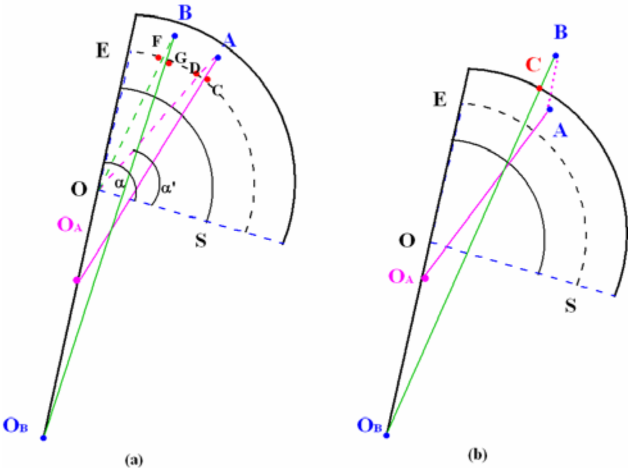

Figure 5 Virtual radius mapping strategies: (a) direct application, (b) boundary projection

4.2 Mapping Strategy for Boundary Condition 2: Edges

It is unavoidable that some users will hit the edges of the road. In normal car dynamics, hitting an obstacle such as a concrete block at the side of the road introduces discontinuities in the speed, acceleration, and angular velocity. We describe here our strategy for mapping the car position so as to preserve tempo smoothness in such conditions.

The method extends the virtual radius mapping (VRM) strategy from the previous section to handle road boundary conditions by assessing virtual position of the car and projecting it back to the edge of the road. Consider Figure 5(b). Suppose the car is first at point *A* at time *t*. If there were no obstacle and the car could go off the road, then it would reach point *B* at time $t + \Delta t$. Since we cannot allow the car to go off road, we will have to map the car position to a point on the side of the road instead of *B*. Using Equation (5) for the radius and the origin as defined in the previous section, O_B in Figure 5(b), we project the car back to the edge of the road, point *C*. We perform this calculation iteratively until the car moves away from the side of the road.

5. PROOF OF SMOOTHNESS

In this section, we will show that the VRM strategies, for both road interior and edge conditions, outlined in the previous sections will guarantee tempo smoothness in expressive rendering of the piece.

We will prove tempo smoothness by showing that the acceleration is guaranteed to vary smoothly over time, i.e. that the change in the acceleration is bounded by a small number for small Δt.

Figure 6 shows the relationship between the music's acceleration, a, and the car's lateral and longitudinal acceleration a_x and a_y. Suppose the curve starts at S, and ends at E, and has a radius R. The car is at point A, with the accelerations a_x and a_y. The virtual origin is O' and the projection point on the central line is B. We can see that BC tangent to the curve at point B, and indicates the direction of the music's acceleration. Let the angle between the car's direction and the tangent direction BC be β. The music's acceleration is given by:

$$a(t) = c(t)(a_y(t)\cos\beta(t) - a_x(t)\sin\beta(t)), \tag{6}$$

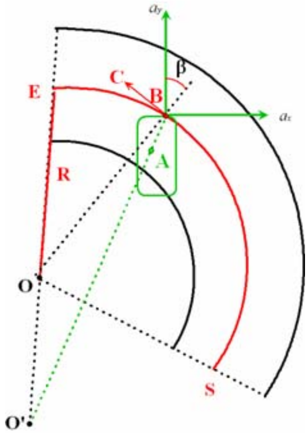

Figure 6 Acceleration of the music

where $c(t)$ is the mapping ratio for the virtual origin O', $O'B/O'A$.

ESP is a discrete system with pre-specified sampling rates. Suppose Δt is small compared to the smallest sampling period. At time t, we will use the dynamics at time $t-1$ to calculate that for time t and $t + \Delta t$. Suppose the user input is reasonably smooth, i.e., that it satisfies the following conditions: $|\delta(t+\Delta t) - \delta(t)| \le \varepsilon_1$, and $|F(t+\Delta t) - F(t)| \le \varepsilon_2$. The goal is to show that when Δt, ε_1 and ε_2 tend to 0, then $|a(t+\Delta t) - a(t)|$ also tends to 0.

By definition,

$$\begin{aligned}
&\left|a(t+\Delta t) - a(t)\right| \\
&=| c(t+\Delta t)[a_y(t+\Delta t)\cos\beta(t+\Delta t) - a_x(t+\Delta t)\sin\beta(t+\Delta t))] \\
&\qquad -c(t)[(a_y(t)\cos\beta(t) - a_x(t)\sin\beta(t)] | \\
&\le \left|c(t)[a_y(t+\Delta t)\cos\beta(t+\Delta t) - a_y(t)\cos\beta(t)]\right| \\
&\quad + \left|c(t)[a_x(t+\Delta t)\sin\beta(t+\Delta t)] - a_x(t)\sin\beta(t)]\right| \\
&\quad + |[c(t+\Delta t) - c(t)][a_y(t+\Delta t)\cos\beta(t+\Delta t) \\
&\qquad -a_x(t+\Delta t)\sin\beta(t+\Delta t)]| \\
&\le \left|c(t)\right|\left|a_y(t) - 2\sin\frac{\beta(t+\Delta t)+\beta(t)}{2}\sin\frac{\beta(t+\Delta t)-\beta(t)}{2}\right| \\
&\quad + \left|c(t)\right|\left|\cos\beta(t+\Delta t)[a_y(t+\Delta t) - a_y(t)]\right| \\
&\quad + \left|c(t)\right|\left|a_x(t)2\cos\frac{\beta(t+\Delta t)+\beta(t)}{2}\sin\frac{\beta(t+\Delta t)-\beta(t)}{2}\right| \\
&\quad + \left|c(t)\right|\left|\sin\beta(t+\Delta t)[a_x(t+\Delta t) - a_x(t)]\right| \\
&\quad + |[c(t+\Delta t) - c(t)][a_y(t+\Delta t)\cos\beta(t+\Delta t) \\
&\qquad -a_x(t+\Delta t)\sin\beta(t+\Delta t)]|
\end{aligned} \tag{7}$$

From the above result, we can check each of the four parts in the resulting equation: $|a_y(t+\Delta t) - a_y(t)|$, $|a_x(t+\Delta t) - a_x(t)|$, $|\beta(t+\Delta t) - \beta(t)|$, and $|c(t+\Delta t) - c(t)|$. If they all tend to 0 when Δt, ε_1 and ε_2 tend to 0, we will have shown that $|a(t+\Delta t) - a(t)|$ also tends to 0. By following this argument, we have proven that smoothness is guaranteed under non-boundary situations, and that it is also preserved under boundary conditions, when we employ the mapping strategies described in the previous sections. The algebraic proofs are too lengthy to include in this paper, and will be provided upon request.

6. EXPERIMENT AND RESULTS

As discussed, our claim is that the mapping strategies proposed will remove discontinuities in the tempo trajectory. In this section, we will demonstrate empirically that this is true. We use Brahms' *Hungarian Rhapsody No. 5* as the test piece for the experiments. In the first 2 experiments, we only play a part of the test piece. This part consists of 96 beats; the excerpt is divided into 12 sections, each corresponding to one segment of the road. The third experiment considers three performances of the entire piece, consisting of 252 beats and 37 road sections.

In ESP, we can readily obtain the instantaneous tempo at every time unit sampled. Hence, we plot the tempo trajectories for analysis. In all the graphs, the horizontal axis represents time. The sampling interval is 0.01 seconds, i.e. 100 sample points per second. The vertical axis represents the tempo in beats per second. We use the value of the objective function in the Minimum Jerk Model [[8]], S, to evaluate the smoothness of the tempo curve.

$$J(i) = a(t) - a(t-1) \tag{8}$$

$$S = \sqrt{\frac{1}{n}\sum_{i=1}^{n} J(i)^2} \tag{9}$$

where S is in beats/second^3.

6.1 Experiment 1: Smoothness under Boundary Condition 1

We first show the difference between the tempo trajectories created using the CLP method, and that using the VRM strategies. In this experiment, a user drives through the piece once, receiving auditory feedback generated using the improved mapping. The user never hits the side of the road in this experiment. We simulate what the output would have been under the CLP mapping method. The tempo for each rendering is plotted in Figure 7, the one above shows the tempo graph for the performance rendered using the VRM strategies, while the one below shows that for the CLP mapping strategy. By inspection, we see that the lower graph contains numerous discontinuities, but the upper one is always smooth. The S values calculated using Equation (8) and (9) concur with the visual analysis: $S_1 = 4.57$ for the improved strategies, and $S_{CLP} = 40.79$ for the CLP strategy.

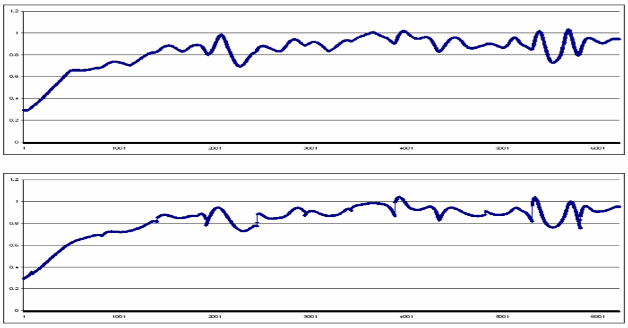

Figure 7 Tempo graphs under virtual radius mapping (top) and centerline mapping (bottom)

6.2 Experiment 2: Smoothness under Boundary Condition 2

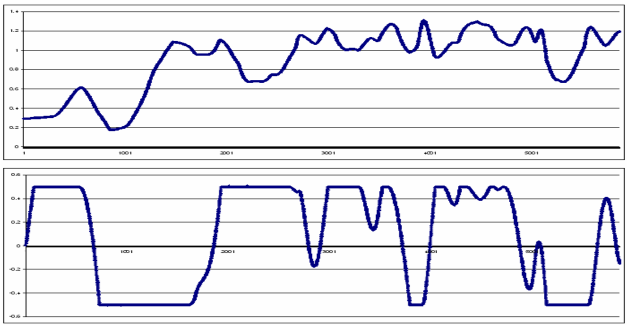

Figure 8 Tempo graph (top) and distance-from-centerline (bottom) for virtual radius mapping under extreme edge-hitting conditions

Next, we demonstrate the effectiveness of the VRM strategy when the car hits the edge of the road. In this experiment, we show the tempo curve for an extremely erratic driver who is instructed to hit the edge of the road many times. The results are shown below in Figure 8. The upper graph in Figure 8 shows the output tempo, while the lower one shows the car's distance from the centerline. A positive value means that the car is on to the right of the centerline, from the driver's viewpoint, and a negative value means that the car is to the left of the centerline. From the distance graph in Figure 8, we can see that the car hits the edge a total of 11 times, because the distance from the centerline reaches the maximum or the minimum value 11 times. In spite of this erratic driving behavior, we can see that the tempo of the music remains smooth. Although there are some sharper changes in the tempo, there are no observable discontinuous points.

6.3 Experiment 3: Performance of the whole piece

The goal of experiment 3 is to show that the system is robust under different conditions of use. One very experienced driver performed the piece three times: the first time driving normally with the proper expressions intended for the piece (results shown in Figure 9, length = 153 seconds), the second time very fast and aggressively (Figure 10, length = 120 seconds), and the third time extremely erratically (Figure 11, length = 125 seconds), hitting the sides of the road numerous times. As before, the chart above shows the tempo graph, while the one below gives the distance from the central line. Visual inspection shows that all performances are reasonably smooth. Even for the most erratic driving style documented in Figure 11, the tempo graph does not have any observable discontinuities.

For each driving style, we calculate the tempo mean and variation, the smoothness value S, and the driving stability as given by the variation of the distance from the centerline. The results are shown in Table 1. They show that the mapping strategies proposed in the paper do ensure the smoothness of the output tempo under various conditions. The normal driving style receives the lowest smoothness value, while the erratic driving style gets the highest. It is useful to note that all smoothness values are much lower than that of the CLP mapping results shown in Experiment 1.

Table 1 Performance Evaluation of the 3 users

	S	Mean	Variation	Stability
Normal	9.72	0.89	0.017	0.043
Aggressive	11.57	1.14	0.044	0.073
Erratic	15.95	1.10	0.063	0.147

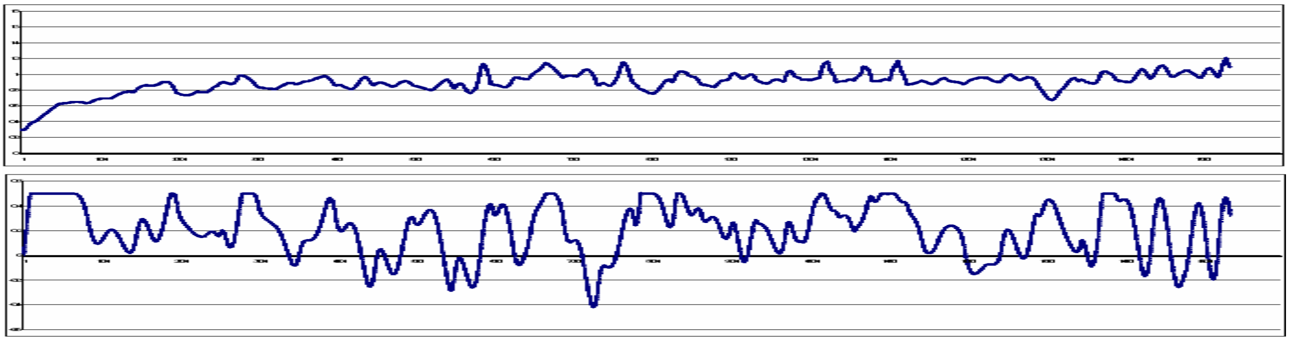

Figure 9 Normal Driving Results under VRM: tempo (top) and distance-from-centerline (bottom) vs. time

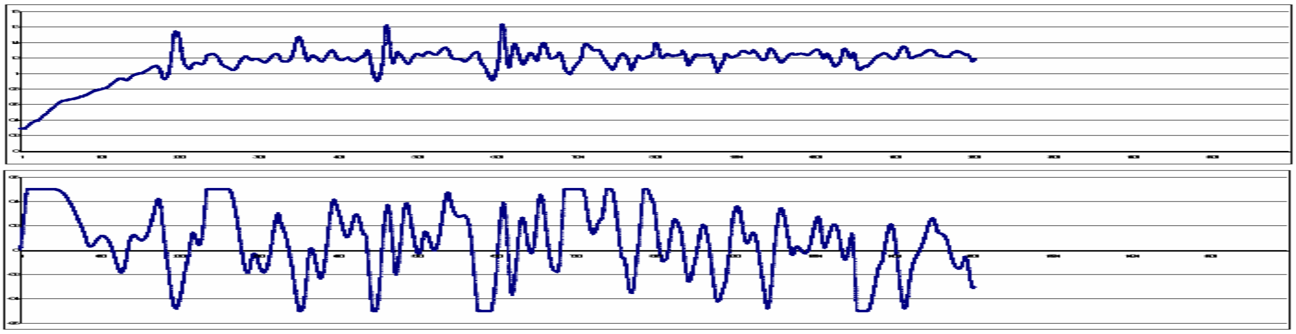

Figure 10 Aggressive Driving Results under VRM: tempo (top) and distance-from-centerline (bottom) vs. time

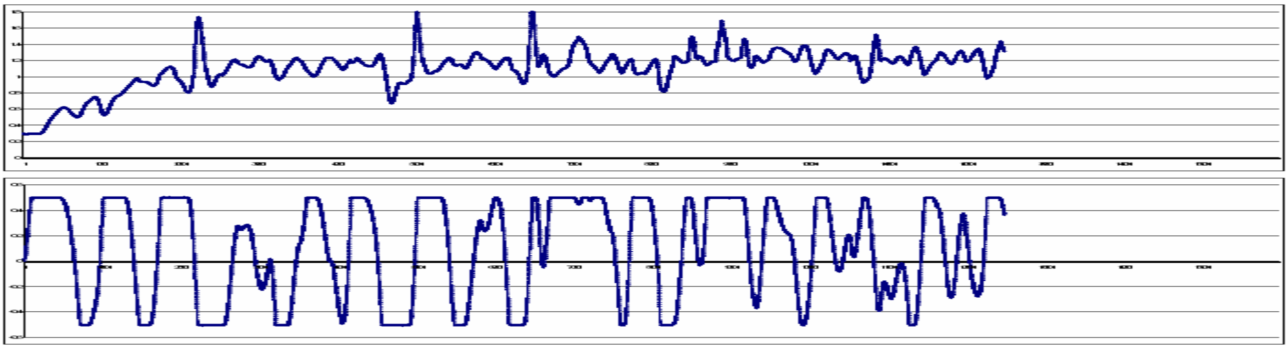

Figure 11 Erratic Driving Results under VRM: tempo (top) and distance-from-centerline (bottom) vs. time

7. CONCLUSION AND DISCUSSION

We began by giving an overview of ESP, a novel interface for creating musical performances by driving through an interactive environment. ESP capitalizes on the driving metaphor for expressive control, thus providing access to high-level decision-making in musical performance for all users that can handle a driving interface. The ESP interface not only allows expert musicians to experiment with expressive choice without having to first master the notes in a piece of music or the instrument, but to also allow non-experts easy access to high-level expressive decisions without having to master an instrument.

A large number of user, musicians and non-musicians of all ages, have tested the ESP interface. The project was showcased at the University of Southern California’s 125th Anniversary celebrations as a featured exhibit in the festival tent on October 6-8, 2005. Pictures from the exhibit, and of visitors using the interface, can be found at [5].

The expert driver in the examples we have presented in this paper is the first author, Jie, who does not have any prior musical background. According to his experience, at the first encounter with a new piece and its corresponding road, the path had great influence on the way he managed the tempo through the piece. As the piece (and road) became more familiar, he was able to drive according to a mental image of how he wished to hear the piece, and the road no longer affected the driving as much. The initial guidance provided by the road of allowed Jie to formulate further interpretations in future traversals of the piece. He now has an excellent grasp of the expressive nuances possible in the piece, and the driving interface allows him to realize various expressive performances without having to first learn to play an instrument.

Because tempo smoothness is a requirement for expert performances, we design mapping strategies in ESP that ensure that the tempo of the rendered performance is smooth under all driving conditions, even the first and possibly awkward encounter with the piece (road). We first described how the road is created from the music score, and we then proposed two mapping strategies that can handle boundary conditions that would otherwise have caused discontinuities in the tempo. The empirical results demonstrate that our mapping strategies do ensure the smooth tempo of the performance rendered, even for the most erratic driving behavior.

Due to the mapping strategies that ensure tempo smoothness under all conditions, any user who creates an expressive performance using ESP can focus totally on experimenting with expressive gestures. The user can fully express his/her own interpretation of the music, without worry of skips in the tempo. Future work will explore other road design methods and interfaces. One option is to let users express their understanding of the music by providing them with building blocks (like that in a train set) for creating their own roads.

8. ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 0347988, and Cooperative Agreement No. EEC-9529152. The work made use of the Integrated Media Systems Center Shared Facilities. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

9. REFERENCES

[1] Car Physics for Games, Marco Monster. http://home.planet.nl/~monstrous/tutcar.html

[2] Cambouropoulos, E., Dixon, S., Goebl, W. and Widmer, G. Human Preferences for Tempo Smoothness. In *Proceedings of the VIII Brazilian Symposium on Computer Music* (Fortaleza, Brazil, 2001)

[3] Chew, Elaine, Alexandre François, Jie Liu and Aaron Yang. ESP: A Driving Interface for Musical Expression Synthesis. *Conference on New Interfaces for Musical Expression*, Vancouver, B.C., Canada, May 26-28, 2005.

[4] Edwards, B., Player Piano History and Evolution. http://www.perfessorbill.com/ragtime5.shtml

[5] Expression Synthesis Project, http://www-rcf.usc.edu/~mucoaco/ESP/

[6] François, A. R. J. A Hybrid Architectural Style for Distributed Parallel Processing of Generic Data Streams. In *Proceedings of the International Conference on Software Engineering (ICSE '04)* (Edinburgh, Scotland, May 2004). pp. 367-376

[7] François A. R. J. Modular Flow Scheduling Middleware (MFSM), 2001. *mfsm.sourceforge.net*

[8] Harris, M. C., Exploring smoothness and discontinuities in human motor behavior with Fourier analysis, *Mathematical Biosciences*, 188 (2004), p99-116.

[9] Hatwal, H. and Mikulcik. E. C. Some Inverse Solutions to an Automobile Path-Tracking Problems with Input Control of Steering and Brakes, *Vehicle System Dynamics 15* (1986), 61-71

[10] Hu, Z., Barnard, E., Smoothness Analysis For Trajectory Features, *1997 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'97)*, Volume 2 p. 979

[11] Katayose, H., and Keita, O. Using an Expressive Performance Template in a Music Conducting Interface. In *Proceedings of the 2004 Conference on New Instruments for Musical Expression (NIME'2004)* (Hamamatsu, Japan, June 3-5, 2004)

[12] Krylow, M. Andrew, Rymer, Zev, W., Role of Intrinsic Muscle Properties in Producing Smooth Movements, *IEEE Transactions on Biomedical Engineering*, Vol. 44, No. 2, February. 1997

[13] Lee, E., Nakra, M.T., and Borchers, J. You're The Conductor: A Realistic Interactive Conducting System for Children. In *Proceedings of the 2004 Conference on New Instruments for Musical Expression (NIME'2004)* (Hamamatsu, Japan, June 3-5, 2004)

[14] Modler, P., Myatt, T., and Saup, M. An Experimental Set of Hand Gestures for Expressive Control of Musical Parameters in Realtime. In *Proceedings of the 2003 Conference on New Interfaces for Musical Expression (NIME'2003)* (Montreal, Canada).

[15] Murphy, D., Andersen, T.H., Jensen, K. Conducting Audio Files via Computer Vision. In *Proceedings of Gesture Workshop* (Genova, Italy, April 15-17 2003)

[16] Sundberg, J., Friberg, A. and Fryden, L., Music and locomotion. A study of the perception of tones with level envelopes replicating force patterns of walking. *Speech Transmission Laboratory Quarterly Progress and Status Report*, 4/1992, pp. 109-122.

[17] Todd, P. M. N, The kinematics of musical expression. *Journal of the Acoustical Society of America*, March 1995, Vol.97, Issue 3, pp.1940-1949

[18] Todorov, E. and Jordan, I. M., Smoothness Maximization Along a Predefined Path Accurately Predicts the Speed Profiles of Complex Arm Movements, *The Journal of Neurophysiology*, Vol. 80, No. 2, August 1998, pp696-714

[19] Xu, Y., Song, J., Nechyba, C.M., and Yam, Y. Performance evaluation and optimization of human control strategy. *Robotics and Autonomous Systems*, 39(2002), 19-36