ESP: A Driving Interface for Expression Synthesis

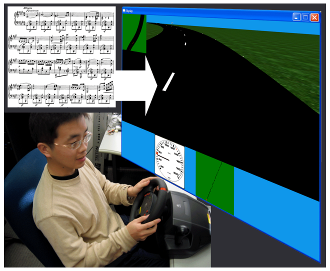
Elaine Chew, Alexandre R. J. François ,
Jie Liu and Aaron Yang

Authors

Introduction

- Goal: to create a driving interface for generating expressive performances interactively in real time from expressionless music files
- Allow non-musicians access to high-level expressive decisions without having to master an instrument
- Allow expert musicians to experiment with expressive choice without having to first master the notes

Introduction



Driving Metaphor

- High-level musical performance is very much like driving a car
  - Score vs. road map
  - Performance vs. a journey on the road
- Harmonixmusic
  - Road
  - http://www.guitarherogame.com/trailer.asp

Driving Metaphor

- Harmonic Driving
  - Visual interface
  - http://gorbet.com/matt/blueharmonic.html
- Tempo smoothness
  - Human prefer smooth tempo transitions
  - Human prefer smooth driving speed

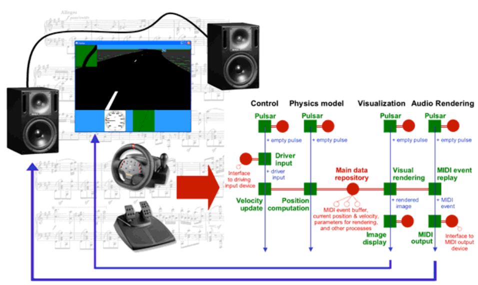
Related Work

- Creating expressive performances
  - The player piano
- Conducting system
  - Radio baton
  - You’re the Conductor
  - Computer vision based baton tracking
  - Digital baton experiment

Related Work

- Tapping-controlled conducting system
- Machine learning model
- Rule-based model

Architectural Design



Car's Physics Model

$$\phi_r = \begin{cases} 0, \phi_f > 0 \\ k_b\phi_f, \phi_f < 0, k_b = 0.34 \end{cases}$$

$$F_{zf} = (mgl_r - (\phi_f + \phi_r)h)/(l_f + l_r)$$

$$F_{zr} = (mgl_f - (\phi_f + \phi_r)h)/(l_f + l_r);$$

$$\alpha_f = \delta - (l_f\omega + v_\xi)/v_\eta = front\_trie\_slip\_angle$$

$$\alpha_r = (l_r\omega - v_\xi)/v_\eta = rear\_trie\_slip\_angle$$

$$\tilde{\alpha}_k = c_k\alpha_k/(\mu F_{zk}), k \in \{f, r\}$$

$$F_{\xi k} = \mu F_{zk}(\tilde{\alpha}_k - (\operatorname{sgn}\delta)\tilde{\alpha}_k^2/3 + \tilde{\alpha}_k^3/27) \times \sqrt{1 - \phi_k^2/(\mu F_{zk})^2 + \phi_k^2/c_k^2}, k \in \{f, r\}$$

$$\dot{\omega} = (l_f\phi_f\delta + l_f F_{\xi r} - l_r F_{\xi r})/I$$

$$\dot{v}_\xi = (\phi_f\delta + F_{\xi f} + F_{\xi r})/m - v_\eta\omega - (\operatorname{sgn} v_\xi)c_D v_\xi^2$$

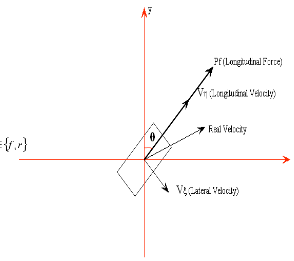
$$\dot{v}_\eta = (\phi_f + \phi_r - F_{\xi f}\delta)/m + v_\xi\omega\,\text{£¨}\operatorname{sgn} v_\eta)c_D v_\eta^2$$

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_\xi \\ v_\eta \\ \omega \end{bmatrix}$$

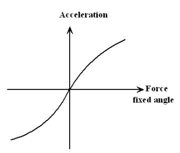
$$X_t = X_{t-1} + \dot{x} * dt$$

$$Y_t = Y_{t-1} + \dot{y} * dt$$

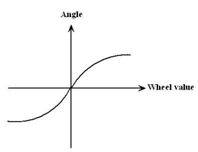
$$\theta_t = \theta_{t-1} + \dot{\theta} * dt$$



Tempo Control



$$a = \lambda_1 F - \lambda_1 v_2 + K_1(\delta, \omega)$$



$$a = -\eta_1\delta - \eta_2 v_2 + K_2(F, \omega)$$

$$T_t = T_{t-1} + a\,dt\,.$$

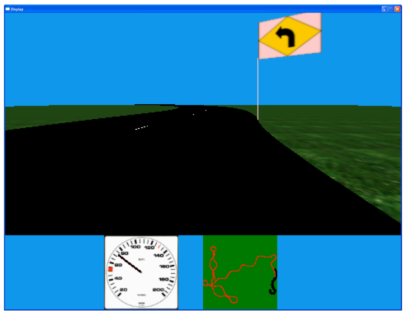
$$D_t = D_{t-1} + T_t dt$$

Acceleration to Loudness

- From Dixon, the rate of tempo change appears to be directly correlated with changes in loudness

$$V_1 = V_0 + a\,\mu$$

Position to Visual Display



Conclusions

- ESP provides a novel and compelling driving metaphor
- ESP allows a user access to high-level interpretative decisions in musical performances
- Not just a driving game
- Not aim to computer-based composing