Music-Guided Aerobatic Flight for F3P Aircraft

Elian van Schalkwijk • Ruihan Jia • Hongze Wang • Prof. Dario Floreano

Motivation

Indoor F3P aerobatics

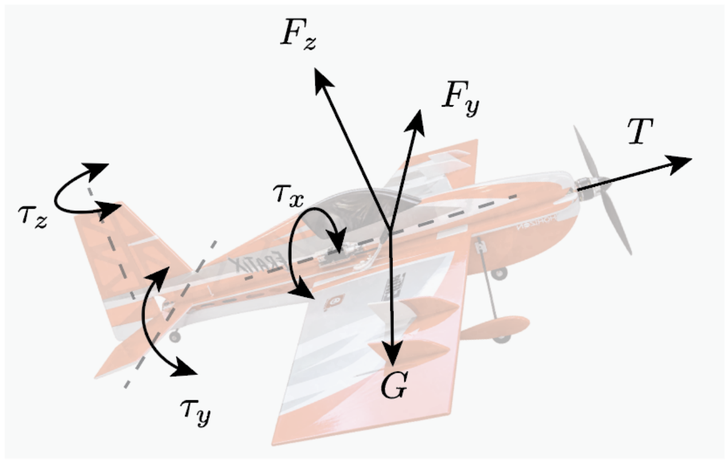
Fixed-wing aircraft, 100 g, 65 cm span

From music to flight

F3P is a form of indoor precision aerobatics where fixed-wing aircraft perform choreographed manoeuvres to music.

We generate aerobatic trajectories from musical beats and train a reinforcement-learning controller to follow them accurately and on time, enabling autonomous music-guided flight.

F3P Dynamics



collective force and torque

$$\mathbf{F}^b = \begin{bmatrix} T \\ 0 \\ 0 \end{bmatrix} - V^2 \begin{bmatrix} k_x \\ k_y \sin\beta \\ k_z \sin\alpha \end{bmatrix} + cV_p^2 \begin{bmatrix} 0 \\ \sin\delta_r \\ -\sin\delta_e \end{bmatrix}$$

$$\boldsymbol{\tau}^b = -V^2 \begin{bmatrix} 0 \\ k_z \ell_H \sin\alpha \\ -k_y \ell_V \sin\beta \end{bmatrix} - V_p^2 \begin{bmatrix} c_a \ell_a & 0 & 0 \\ 0 & c\ell_t & 0 \\ 0 & 0 & c\ell_t \end{bmatrix} \begin{bmatrix} \sin\delta_a \\ \sin\delta_e \\ \sin\delta_r \end{bmatrix}$$

V	airspeed (V_p: propwash)	δ	control surface deflection
α	angle of attack	k	airframe force coefficient
β	sideslip angle	c	control surface force coefficient
T	thrust	ℓ	moment arm

Approach

1. **Beats.** Extract beat times and strengths from the music.
2. **Waypoints.** Convert each beat into a 3D waypoint and turn.
3. **Fitting.** Build a continuous min-snap trajectory.
4. **Following.** Train PPO to track the reference trajectory.

Setup

- Aircraft: 100 g, 65 cm span, 0.5–7 m/s.
- Policy: 50 inputs, 4 actions, 15 177 parameters.
- Training: 1200 iterations × 16 episodes; control at 10 Hz.

Trajectory Fitting

The waypoints are converted into a continuous trajectory using piecewise seventh-order polynomials, with minimum-snap optimisation.

$$u = \frac{t - T_i}{\Delta t_i} \in [0,1], \quad p_i(u) = \sum_{k=0}^{7} c_{i,k}\, u^k$$

$$\min_c \sum_{i=0}^{m-1} \int_{T_i}^{T_{i+1}} \left\| \frac{d^4 x}{dt^4} \right\|^2 dt \;=\; \min_c \; c^\top Q\, c$$

Piecewise polynomial
Eight coefficients per segment and per axis; derivatives are obtained analytically.

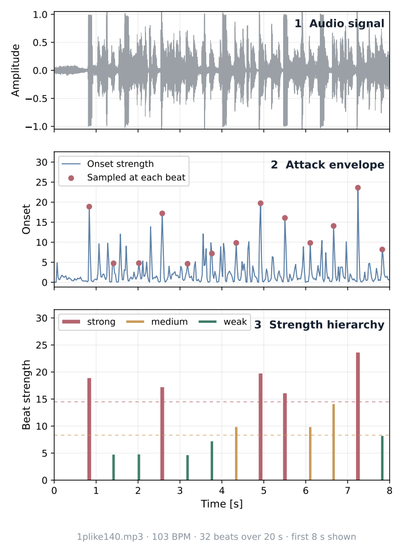
Minimum snap
Minimising the integrated fourth derivative of position favours smooth trajectories while satisfying the waypoint constraints.

Constraints

Position and velocity are imposed at every waypoint, with continuous acceleration at segment junctions, giving a C^2 trajectory. Corner anchors localise the turn around sharp waypoints.

Cube: 22 segments, $f = 0.45$. **Figure of eight:** 16 segments, $f = 0$.

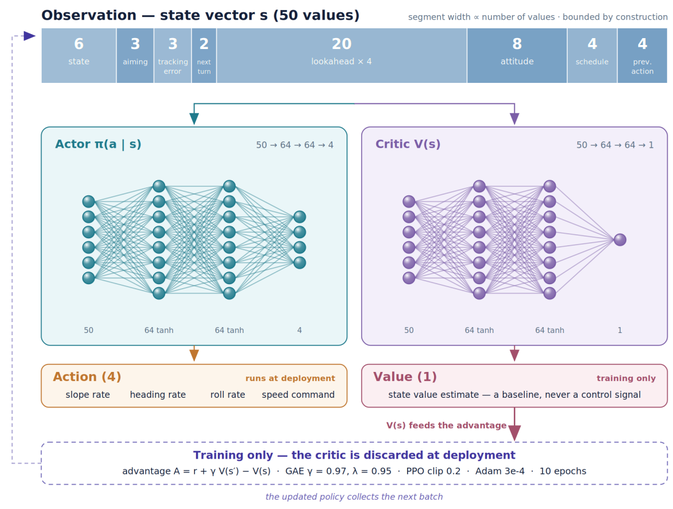
PPO: Neural Network and Reward



1plike140.mp3 · 103 BPM · 32 beats over 20 s · first 8 s shown

Onset strength and beat hierarchy

beats → waypoints → reference



Actor–critic MLP, separate networks

The policy combines the aircraft state, its previous action and the future reference states to predict continuous control commands, constrained by the measured flight envelope to ensure feasible motion.

Reward

$$r_t = \underbrace{r_{\text{beat}}}_{\text{punctuality}} + \underbrace{W_{ref}\, e^{-e_{ref}/1.5}}_{\text{tracking}} + \underbrace{W_{sch}\big(|\ell_{t-1}| - |\ell_t|\big)}_{\text{schedule}} + \underbrace{W_{\mu}\, e^{-|\Delta\mu|}}_{\text{roll}} - \underbrace{W_{dact}\, \|a_t - a_{t-1}\|^2}_{\text{jerkiness}}$$

e_{ref} distance to the reference point, ℓ_t signed schedule error, $\Delta\mu$ roll error, a_t action; the W are fixed weights.

$$r_{\text{beat}} = \sum_{i\,:\,T_i \in\,]t-\delta t,\, t]} W_{TIME} \exp\!\left(-\frac{\|p(T_i) - W_i\|}{D_{TIME}}\right)$$

$$\ell_t = \big(p_t - x_{ref}(t)\big) \cdot \hat{d}_{ref}(t), \qquad \Phi(s) = -|\ell|$$

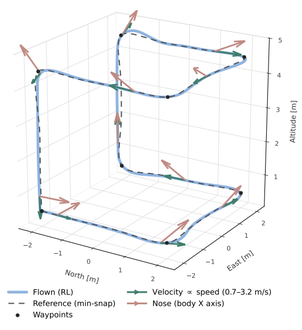
Punctuality
Rewarded at the exact beat instant. $W_{TIME} = 20$, $D_{TIME} = 1\,\text{m}$.

Schedule
The potential difference rewards progress along the reference trajectory at every step.

Trajectory Generation and Following

From beats to waypoints — rule-based generation

Each beat becomes a 3D waypoint. Beat strength sets the turn magnitude and plane, while the beat interval sets the segment timing.



Cube — 9 beats: waypoints, reference and flown path

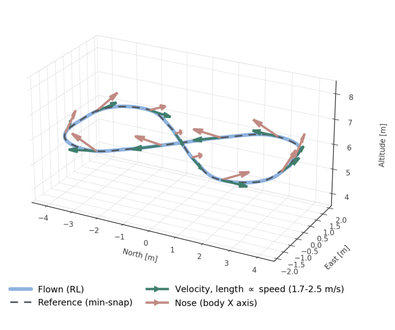
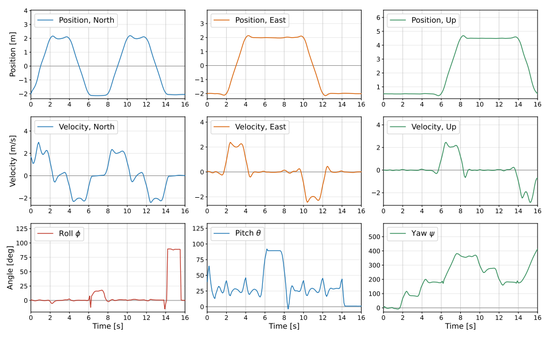
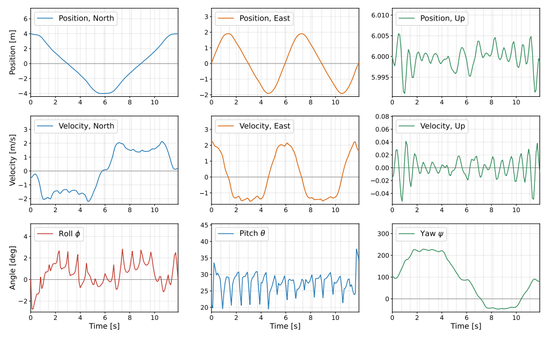


Figure of eight: waypoints, reference and flown path



Flight data — cube



Flight data — figure of eight

Following — PPO from scratch

The policy outputs three rate commands and a speed command, trained from scratch for 1200 iterations.