MoDRE 2026 • MODEL-DRIVEN REQUIREMENTS ENGINEERING WORKSHOP

A Model for Mediating Multi-Modal Human Intent into Safe Maneuvers for UAVs

Sofia Nelson
University of Notre Dame
Notre Dame, IN, USA
snelso24@nd.edu

Dalal Alrajeh
Imperial College London
London, UK
dalal.alrajeh@ic.ac.uk

Pedro Antonio Alarcon Granadeno
University of Notre Dame
Notre Dame, IN, USA
palarcon@nd.edu

Jane Cleland-Huang
University of Notre Dame
Notre Dame, IN, USA
janehuang@nd.edu

MOTIVATION

Human–drone Interaction Made Simple

- Small UAVs act as collaborative teammates under human supervision
- Operators occasionally step into the loop with direct maneuver commands
- Used when human judgment must temporarily override autonomy

Speech

"Go forward"

Gesture

arm poses

GUI

on-screen buttons

Related Work

- Assumes recognized commands are executed directly
- Generally focuses on human comfort, separation distance, and safety envelopes
- Does not evaluate commands before executing them

Fly into terrain

descend below safe ground clearance

Collide with another UAV

violate inter-drone separation in the air

Collide with the operator or another person

collide with a person on the ground

THE SOLUTION

Treat operator inputs as **maneuver requests;** not actuation commands

PAIR

Pair the drone to a specific human operator

REQUEST

Interpret intent with a confidence score

EVALUATE

Validate against terrain, separation & envelope

RESPOND

Constrain, reject, or execute — then monitor

THE MODEL

M³R: the UAV is always in one of six states

NormalAutonomy

The drone performs its mission under normal autonomous control

ManeuverResponseArmed

After pairing with a human operator, he drone is eligible to receive maneuver commands

ManeuverResponseActive

A maneuver request has been accepted and translated into a bounded motion primitive

Blocked

A requested maneuver has been rejected because it would violate one or more safety constraints.

EmergencyHold

The UAV has interrupted execution and transitioned to a safe hover or hold state.

ReturnToAnchor

The drone returns to a safe anchor point after a maneuver or interruption

M^3R Framework

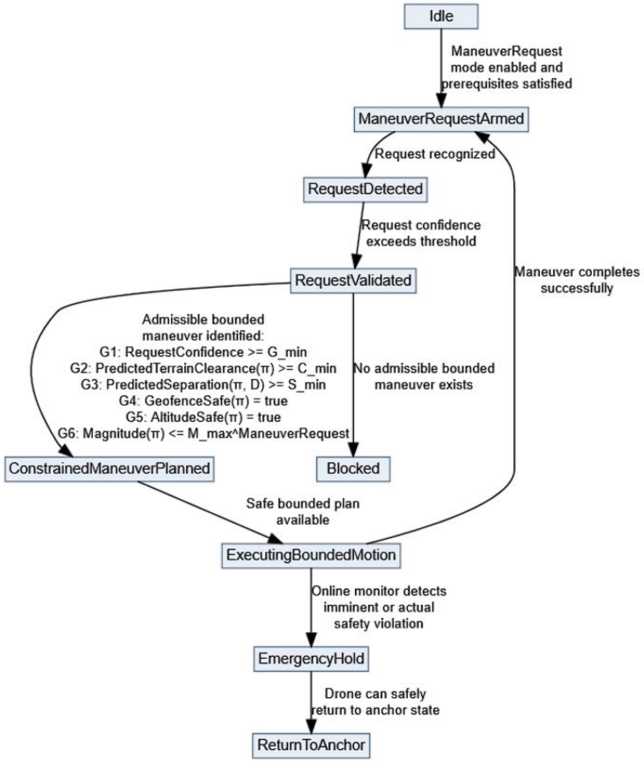
M^3R Model

Formal Synthesis

- Treat the model as a formal specification
- Automatically synthesize a controller (e.g. GR(1))
- Correct-by-construction; formally prove safety

Traditional SE

- Use the model to guide a traditional SE process
- Human-built; analyzable, testable, traceable



THE PIPELINE

The Pairing Pipeline

One operator ↔ one UAV

1

Operator sends a pairing request

›

2

UAV rotates to face the operator

›

3

UAV confirms the pairing

THE PIPELINE

Operator view vs drone view

DIRECTIONS ARE OPERATOR-RELATIVE

“Move right” ⟶ the UAV moves to ***its*** left.
The UAV holds an operator-facing orientation, so commands are read from the human’s perspective!

Requests map to bounded motion primitives

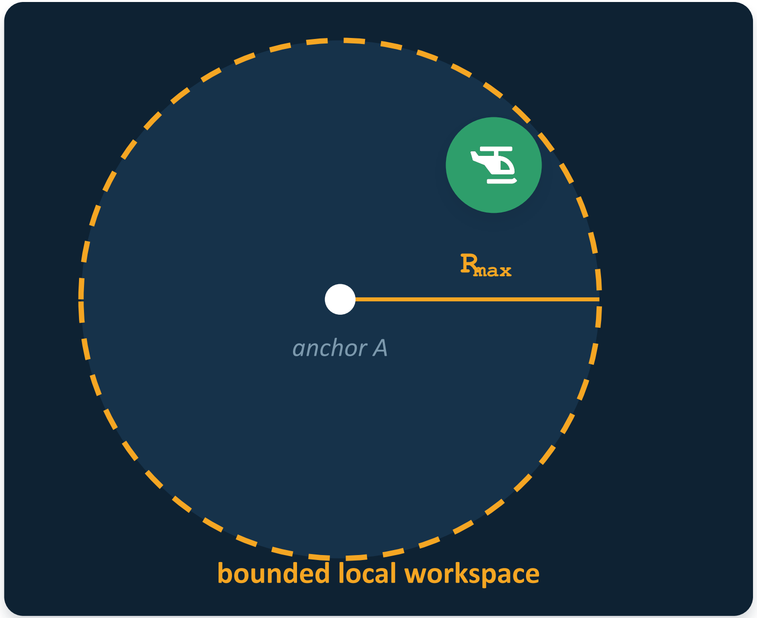
Request	Primitive	Example Bound
Go left / right	MoveLeft / MoveRight	≤ 1 m lateral
Go forward / back	MoveForward / MoveBackward	≤ 1 m translation
Go up	Ascend	≤ 0.5 m vertical
Go down	Descend	≤ 0.5 m, reduced speed
Turn left / right	ControlledYawScan	360°, returns to operator-facing

EVERY PRIMITIVE INCLUDES

- Intended direction
- Maximum displacement
- Maximum speed
- Maximum duration
- Completion condition
- Safety preconditions
- Abort conditions

A bounded workspace around a dynamic anchor

- On arming, the UAV records an anchor state $\mathbf{A} = (x_A, y_A, z_A, \psi_A)$
- Every subsequent maneuver is confined to a sphere of radius R_{max} around that anchor
- The anchor/boundary both update after every accepted command
- The sphere is further clipped by real-world variables



Verify Safety of Command

1. `intent = Interpret(g)`

   recognized request ⟶ intent

2. `π_cand = Plan(intent, x)`

   intent ⟶ candidate path

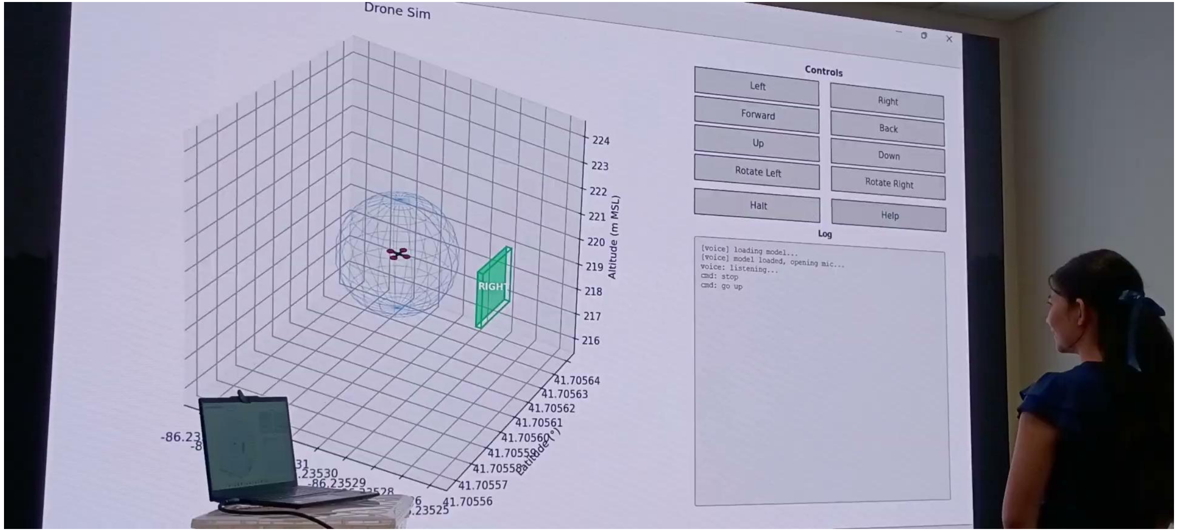
3. `π_safe = Filter(π_cand, T, D, C)`

   candidate ⟶ filtered path

```
Safe(π, x, T, D, C)  :=
      TerrainSafe
  ∧   SeparationSafe
  ∧   EnvelopeSafe
  ∧   ConfidenceAdequate
```

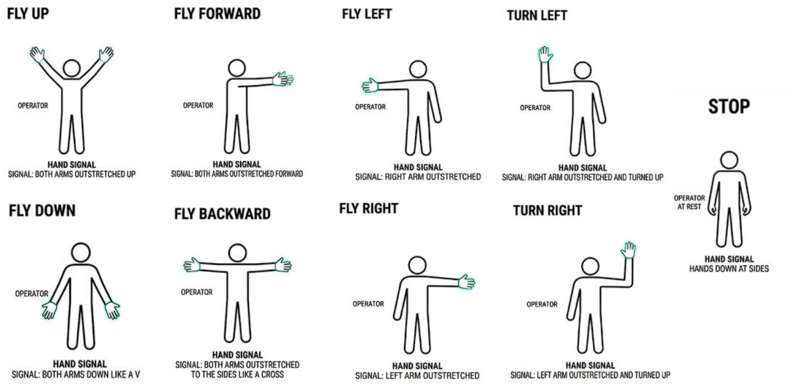
If Safe is false ⟶ the maneuver is reduced, delayed, rejected, or replaced with a safer alternative.

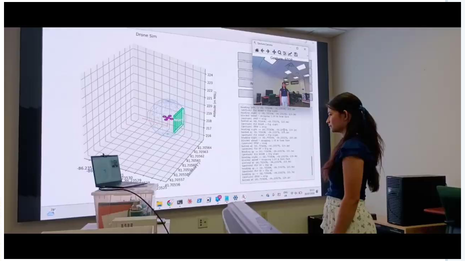
Modality Interfaces: Voice-Based Maneuvers



Voice activated maneuvers are recognized using VOSK (built on Kaldi's speech recognition framework). NLP cleaning is applied to filter out noise, and best match is applied.

Modality Interfaces: Gesture-Based Maneuvers





9 commands via a Visual LLM (GPT-5.0) on the video stream (to be replaced by a YOLO model trained on more data).

FUTURE WORK

What Comes Next

Better CV Model

- Replace VLLM with YOLO
- Train on NOMAD Aerial Dataset

Synthesis & verification

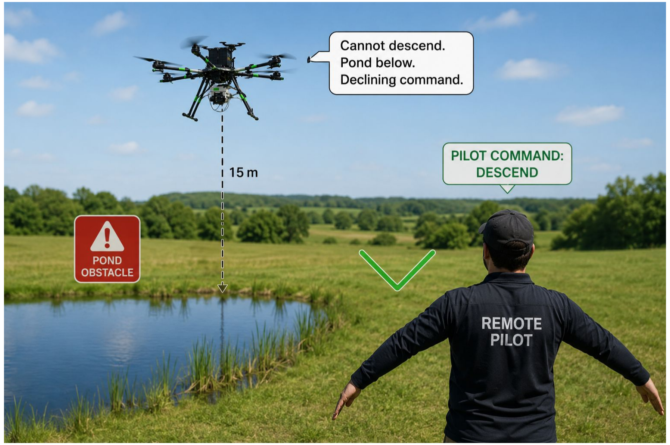
- Synthesize a controller directly from the model
- Formally prove safety properties

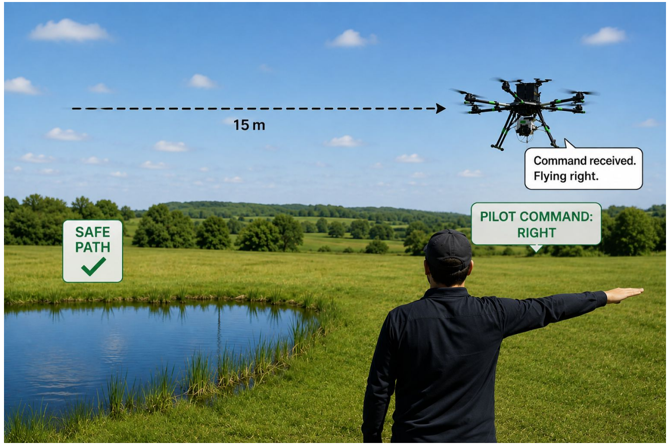
Outdoor field testing

- Physical UAVs outdoors under wind, low visibility, etc.
- Static (terrain, structures) and dynamic (other UAVs) obstacles

Examples


10 m
Cannot fly right
– trees in the way
RPIC
Request: Fly right
Response: Declined
Reason: Obstacle (trees)


Cannot descend.
Pond below.
Declining command.
15 m
PILOT COMMAND:
DESCEND
POND
OBSTACLE
REMOTE
PILOT


15 m
Command received.
Flying right.
PILOT COMMAND:
RIGHT
SAFE
PATH

CONCLUSION

Safe multimodal UAV interaction cannot rest on recognition alone

- Speech, gesture & GUI inputs are mediated as bounded requests; evaluated against terrain, separation, confidence & envelope before execution
- Safety checks and state machine enable verification and future reactive synthesis
- An initial prototype shows inputs can be safely constrained; halting or limiting maneuvers when a violation is detected

Thank You!

Sofia Nelson snelso24@nd.edu