Image Based Visual Control of Aerial Robotic Vehicles

Robert Mahony
College of Engineering
and Computer Science
Australian National University

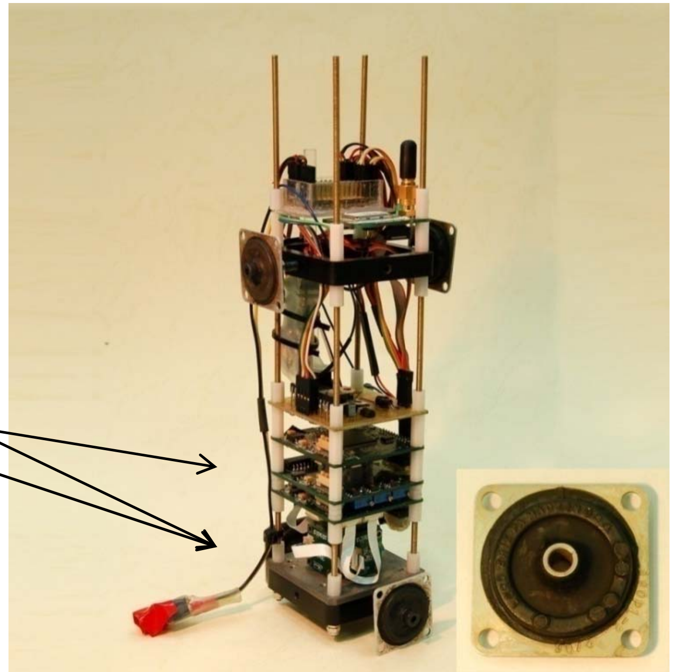
AETOS conference, Bordeaux,
25 September 2012.

An aerial robot is an unpiloted aerial vehicle equipped with sensors that enable it to understand and adapt to changes in its surrounding conditions

(my definition)

- According to this definition, many existing unmanned aerial vehicles (UAVs) are not aerial robots.
  - Pre-programmed tracking of GPS way points does not allow for adapting to changing local environment.
- Aerial robots require *exteroreceptive* sensor systems capable of providing local information for complex 3D-environments.

Avionics stack from ANU X4-flyer



Proprioceptive sensors

Inertial Measurement Unit
- Accelerometers
- Gyrometers
- Magnetometers

Global positing system (GPS)
(for outdoors applications)

Extereoreceptive Sensors

Vision
Laser – IR range
Structured light

Physical noise isolation plus suitable analogue and digital post filtering.

- Vision systems are light (as little as a few grams)
- Vision is passive requiring low power and producing no signature emissions.
- Vision sensors are one of the most information rich exteroreceptive sensors that exist.
- Vision sensors provide so much information that it is possible (in principle) to understand complex 3D-environments using vision alone (structure from motion)

Caveat: *Effective post processing of vision requires significant computational resources.*

Classical control theory provides a standard approach to regulation problems:

1. Model the dynamics of the system.
    - Represent the dynamics in terms of a ***minimal state***.
2. Represent the task in terms of a state error.
3. Design a control algorithm to drive the state error to zero.

---

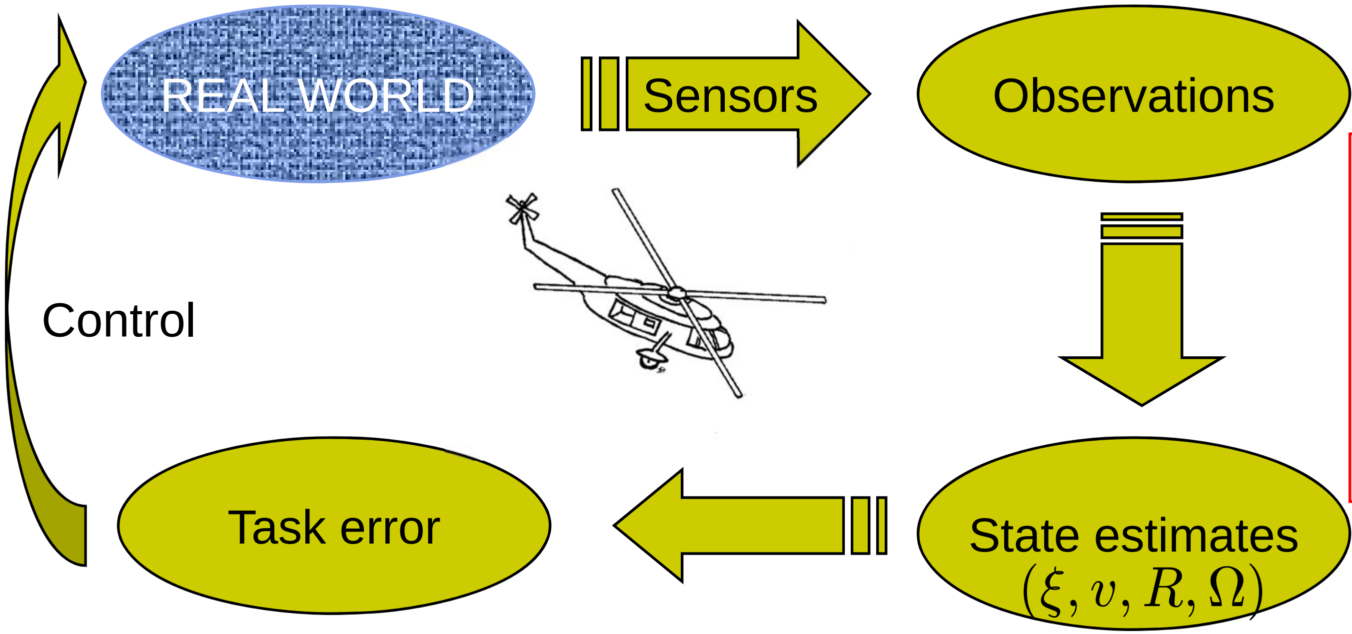
4. Measure as much as you can.
5. Estimate the system state on-line.

---

6. Input the state estimate into the control algorithm to close the loop.

Conceptual framework of state-based control

Australian National University



The mapping from observation to state estimate is
- non-linear
- over-determined
- ill-conditioned

Task error is often naturally conditioned relative to proximity to environment!

Easy to represent in terms of sensor measurements.

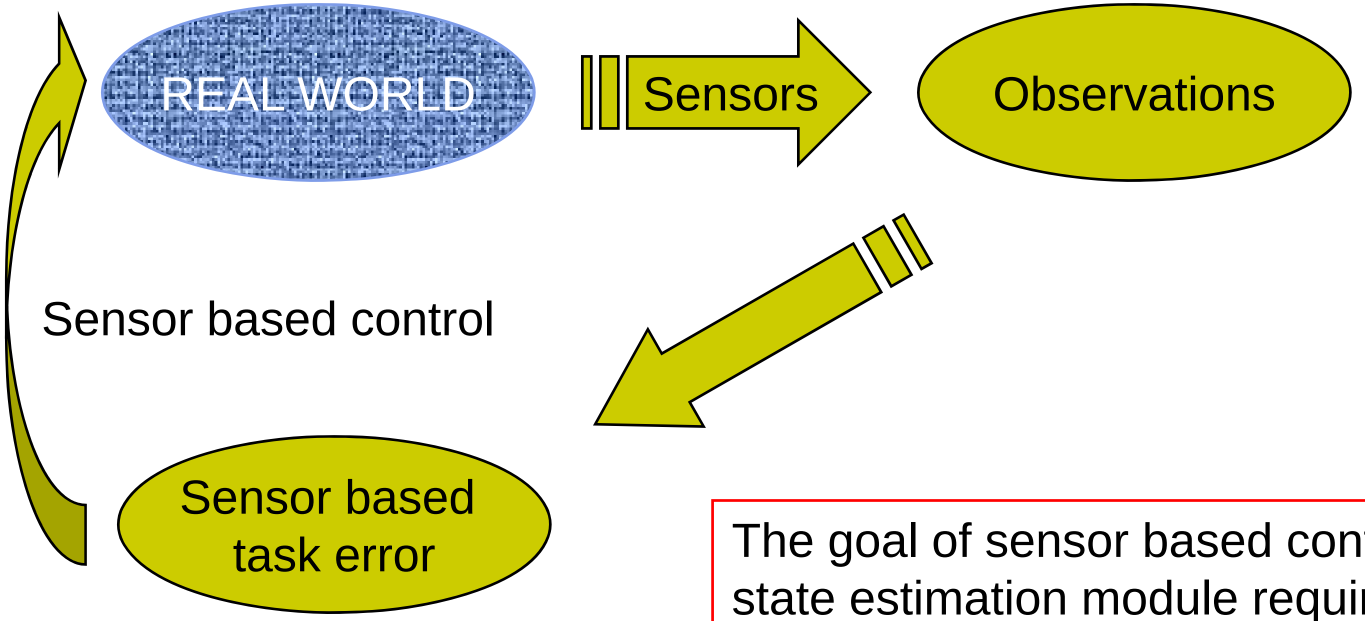
Computing a state estimate from the observations requires:
- Model of the environment (SLAM)
- Model of the system dynamics

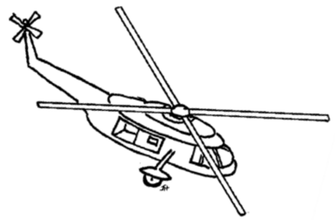
Estimates tends to be ill-conditioned when the vehicle is distant from local features.

Sensor based control is a paradigm that is only subtly different from the classical approach.

1. Model the dynamics of the system
   - Use this model to determine the dynamic response of the *sensor signals* based on *general structure assumptions about the environment*.
2. Represent the task in terms of a sensor error
3. Design a control algorithm to drive the sensor error to zero based on analysis of the sensor and system dynamics
4. Input the sensor measurements into the control algorithm to close the loop.

Conceptual framework for sensor based control





The goal of sensor based control is to remove the state estimation module required for a classical systems and control approach.

It is particularly effective in situations in which state estimation is difficult or ill posed, and the task can be directly represented in terms of measurements.

Extract a set of point features from the observed image in real time.

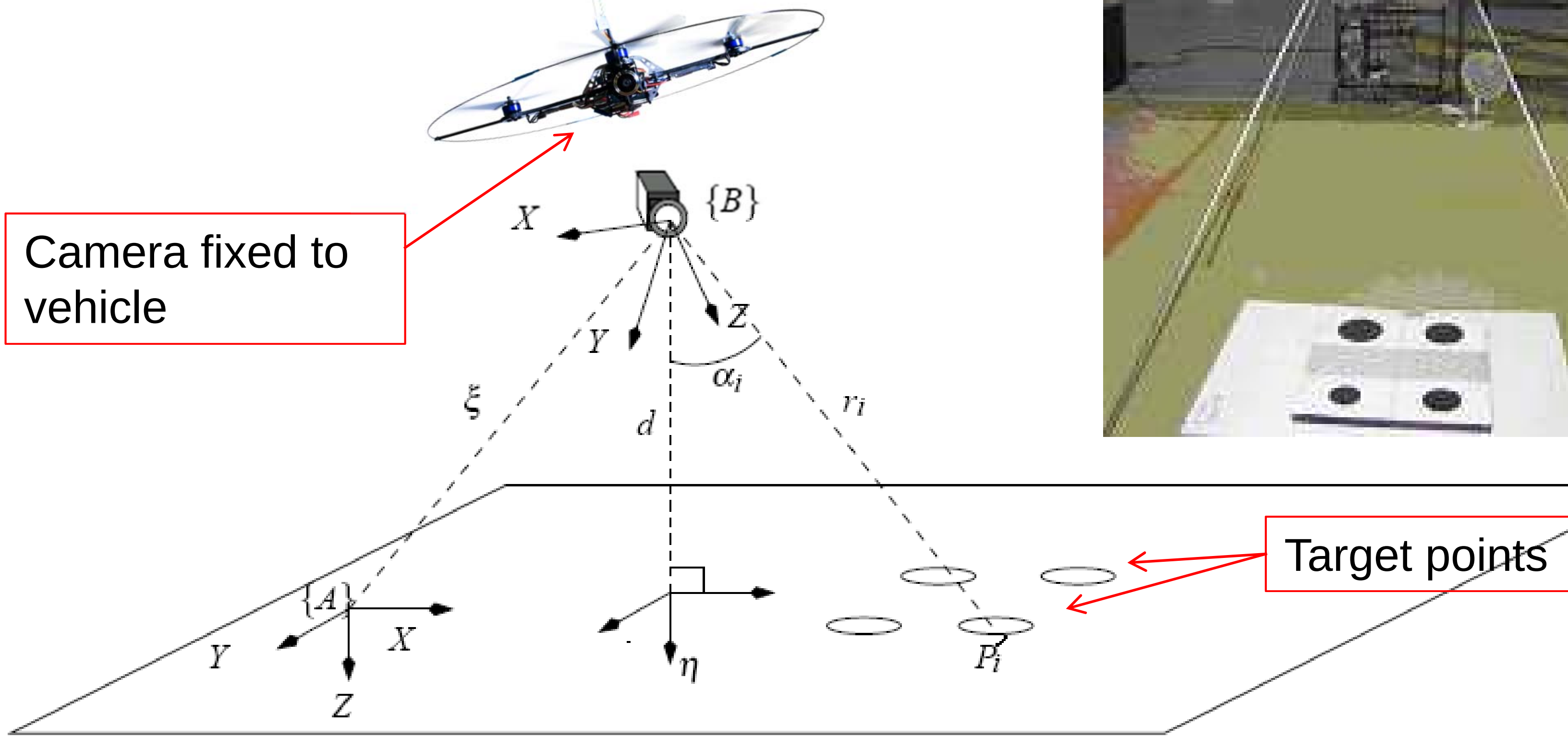


Image points s_i, desired image points s_i^*, image error s_i^*, leading to image kinematics

$$\dot{s} = L(s_i, z_i, q)\dot{q}, \qquad \dot{q} = -L^{\dagger}(s_i, z_i, q)(s_i - s_i^*)$$

Extending to image dynamics appears hopeless

$$\ddot{\tilde{s}} = \frac{d}{dt} L(s_, z_i, q)\dot{\tilde{q}} + L(s_, z_i, q)\ddot{\tilde{q}}$$

There are two reasons for this difficulty

- The image coordinates have not been given any physical interpretation.
- Classical image based visual servo control is based on linearising control design.

A port Hamiltonian dynamic systems model approach

Australian National University

Consider a classical Euler–Lagrange dynamical system with generalised coordinates *q*

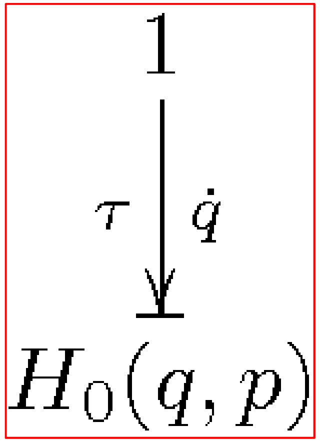
$$M(q)\ddot{q} = -C(q,\dot{q})\dot{q} - \frac{\partial U}{\partial q}(q) + \tau$$

$$p = M(q)\dot{q}$$

The Hamiltonian associated with this is

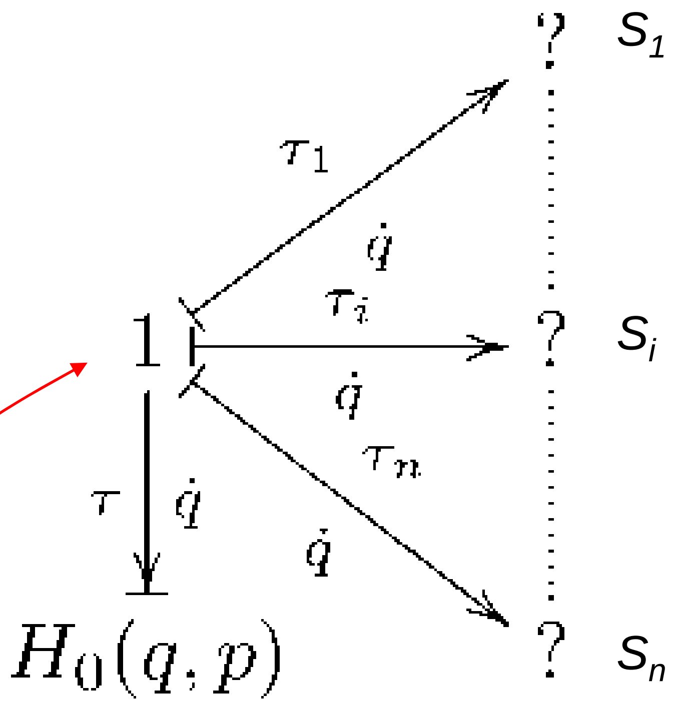
$$H_0(q,p) := \frac{1}{2}p^\top M^{-1}(q)p + U(q),$$

The power bond is



$$\frac{d}{dt}H_0(q,p) = \langle \tau \,|\, \dot{q} \rangle = \tau^\top \dot{q}$$

Initially we consider each image feature s_i as a modular bond interconnection



Thus, each image feature s_i will result in a separate bond attached to the system 1-junction coupling separate effort variables τ_i to the Hamiltonian dynamics

$$0 = \tau + \tau_1 + \cdots + \tau_n$$

To interface the ith image bond to the image feature s_i is a on-linear modulated transformer

$$\xrightarrow[\dot{q}]{\tau_i} \mathrm{MTF} \xrightarrow[\dot{s}_i]{e_i}$$

The transformer relationship is defined by the relationship between image flow and the generalized velocity

$$\dot{s}_i = L_i(s_i, z_i, q)\dot{q}, \quad \Longleftrightarrow \quad \tau_i = L_i(s_i, z_i, q)^\top e_i$$

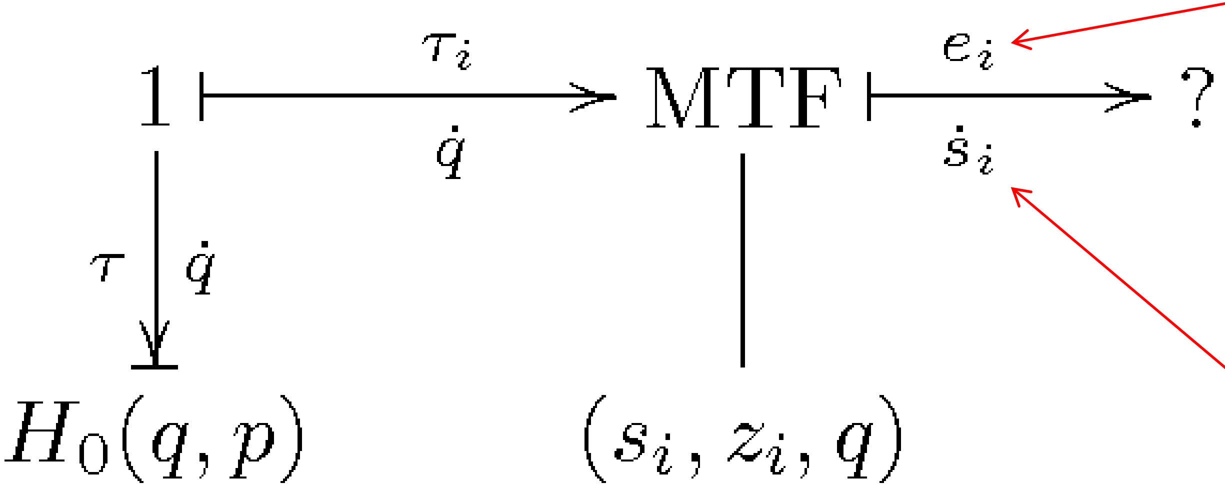
Conservation of power relates the effort e_i to τ_i.

Image bond

Australian National University

$$\dot{s}_i = L_i(s_i, z_i, q)\dot{q}$$

Modulated transformer



The effort e_i will be determined by nature of the element used to terminate the bond.

The flow $\dot{s}_i$ is the image flow of the observed feature.

Modulated transformer depends on signals (s_i, z_i, q)

An *image Hamiltonian* H_i is a spring like storage element defined in the image space

$$H_i(s_i) := \frac{1}{2} k_i ||s_i - s_i^*||^2, \qquad k_i > 0.$$

The standard model yields (Hook's law)

$$e_i := \frac{\partial H_i}{\partial s_i} = k_i(s_i - s_i^*)$$

$$\longmapsto \xrightarrow[\dot{s}_i]{e_i = \partial H_i / \partial s_i} H_i(s_i)$$

The power balance equation is

$$\frac{d}{dt} H_i = \left(\frac{\partial H_i}{\partial s_i} \right)^\top \dot{s}_i = \langle e_i \,|\, \dot{s}_i \rangle = k_i (s_i - s_i^*)^\top \dot{s}_i$$

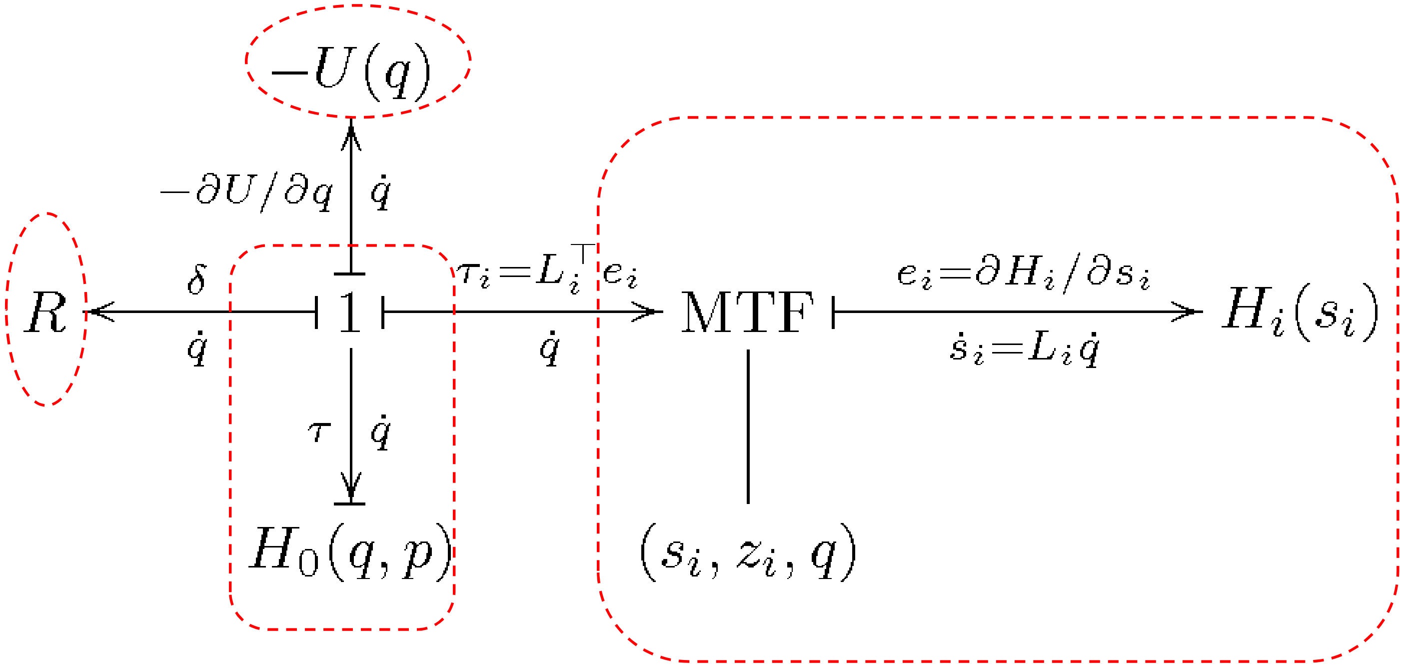
Full modular dynamic IBVS

Australian National University

Compensate for the internal potential of the Hamiltonian H_0

Add damping to obtain stable response.

$$\delta = R\dot{q}$$
$$R > 0$$



System dynamics

$$\tau = -\sum_{i=1}^{n} \tau_i - \delta + \frac{\partial U}{\partial q}(q)$$

Image Hamiltonian used to shape potential for desired regulation point.

System dynamics of closed-loop system is

$$M(q)\ddot{q} = -C(q,\dot{q})\dot{q} - \sum_{i=1}^{n} \tau_i - \delta = -C(q,\dot{q})\dot{q} - \sum_{i=1}^{n} k_i L_i^\top(s_i, z_i, q)(s_i - s_i^*) - R\dot{q}$$

Gain tuning is important to get desirable transients

Note that the image Jacobian is not inverted!

Total energy of closed-loop system

$$\begin{aligned} H(q,p,s_i) &= H_0(q,p) + \sum_{i=1}^{n} H_i(s_i) - U(q) \\ &= \frac{1}{2} p^\top M^{-1}(q) p + \frac{1}{2} \sum_{i=1}^{n} k_i \|s_i - s_i^*\|^2 \end{aligned}$$

Energy evolution

$$\frac{d}{dt} H(q,p,s_i) = -\langle \delta \mid \dot{q} \rangle = -\dot{q}^\top R \dot{q}.$$

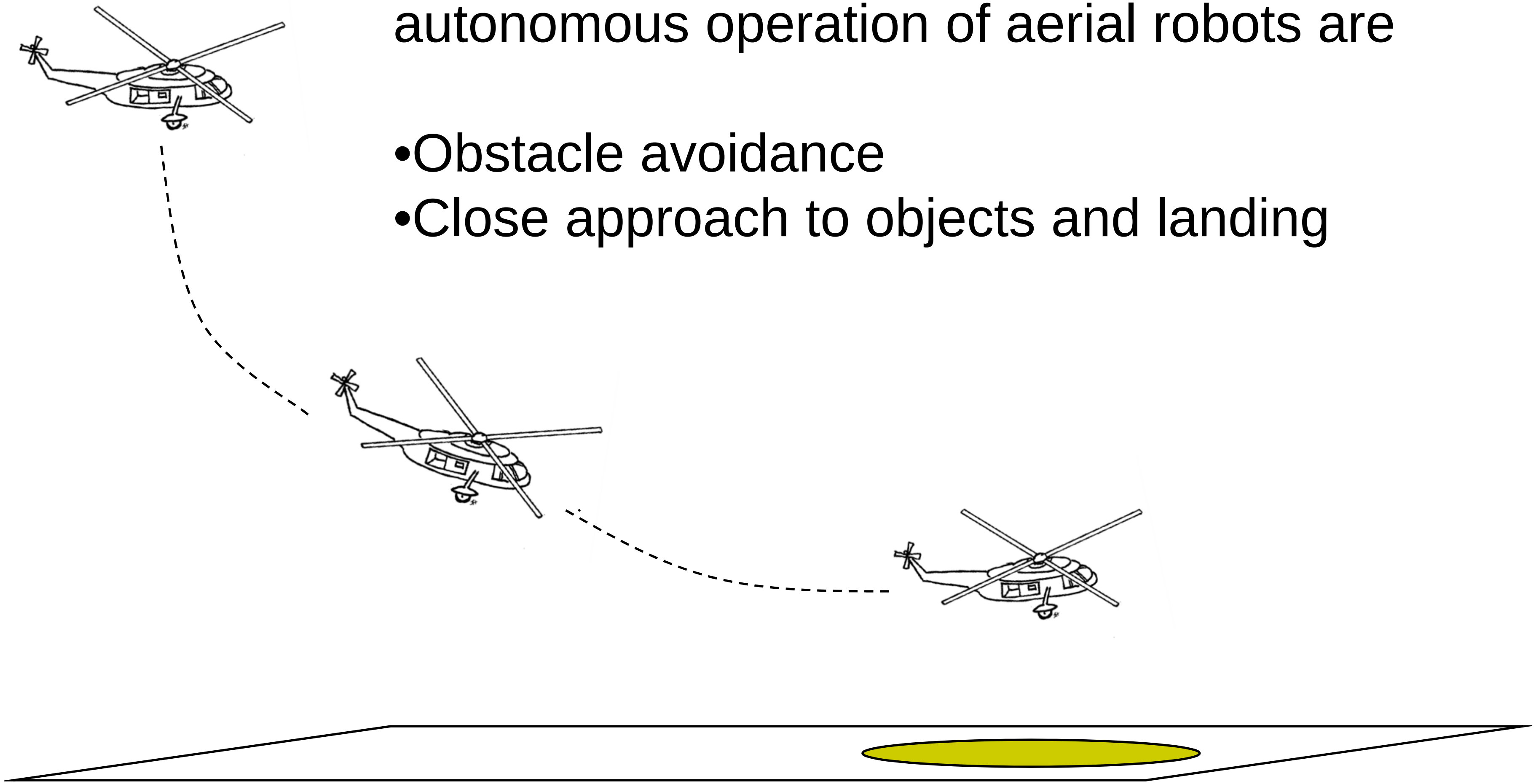
Straightforward to prove stability.

Vision based Landing and obstacle avoidance

Australian National University

Two key tasks that are necessary for autonomous operation of aerial robots are

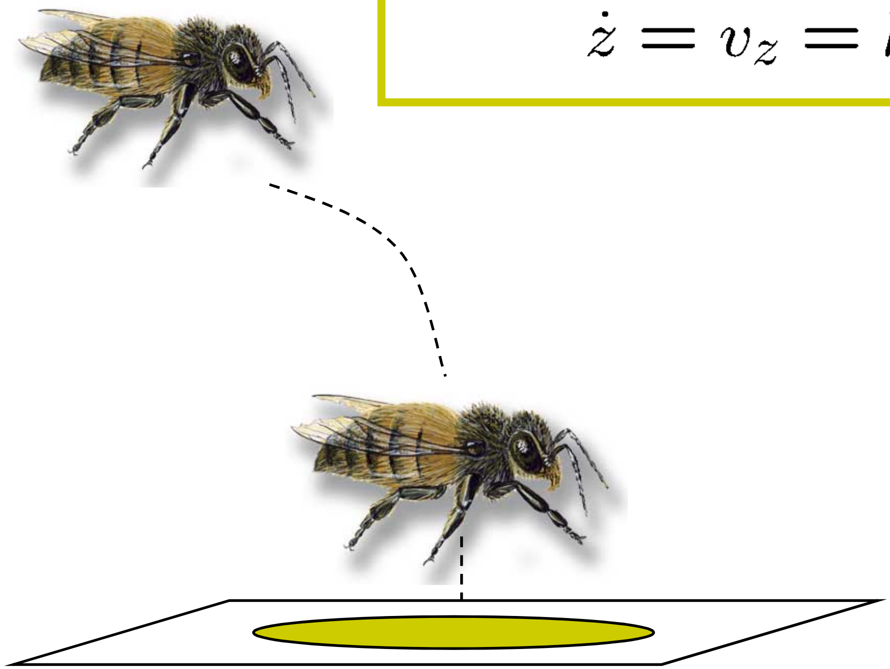
- Obstacle avoidance
- Close approach to objects and landing



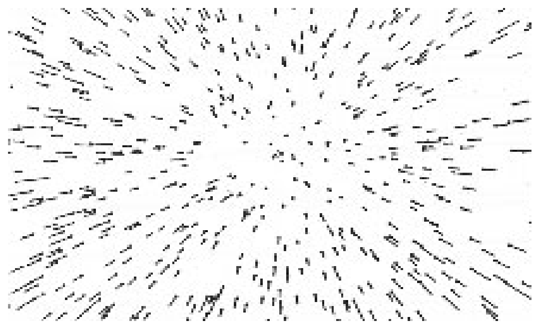
A key insight that can be used in developing a motion cue for such a task comes from the study of biological systems.

A honey bee regulates its thrust in landing approach in proportion to a measure of divergence of the observed optic flow (Srinivasan et al. 2000)

$$\dot{z} = v_z = k\,(\mathrm{div}\Phi - \phi_{\mathrm{ref}})$$



$$\Phi \approx -\frac{v_z}{z}$$



Optical flow field ϕ of textured surface under direct approach

Consider the spherical optical flow for a camera with rigid-body velocity (V, Ω)

$$\begin{aligned}\Phi(\eta) &= \quad \Theta(\eta) \quad + \quad \Psi(\eta), \\ &= (-\Omega \times \eta) + \left(-\frac{1}{\lambda(\eta)} \left(I - \eta\eta^\top \right) V \right).\end{aligned}$$

Spherical optic flow splits into ego-rotation $\Theta(\eta)$ and ego-translation $\Psi(\eta)$.

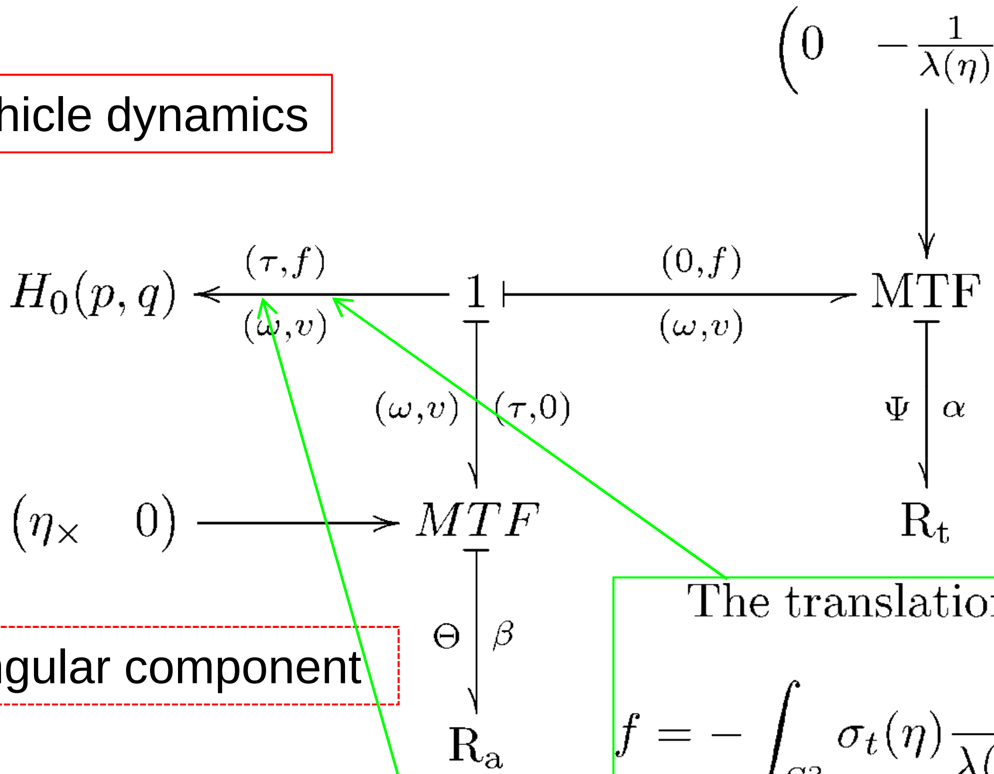
Define a positive definite operator $R_\eta : T_\eta S^2 \to T^*_\eta S^2$.

Applying R_η pointwise to an optic flow field defines an infinitesimal visual effort

$$\begin{aligned}e &: \Gamma(TS^2) \to \Gamma(T^*S^2 \otimes \wedge^2(T^*S^2)), \\ e(\Phi_t(\eta)) &:= \sigma_R(\eta)\Phi_t^\top R_\eta(\cdot)\, d\theta_\eta \wedge d\psi_\eta\end{aligned}$$

Optic flow motion damping

Vehicle dynamics



Translational component

Angular component

The translational force applied by the dissipation

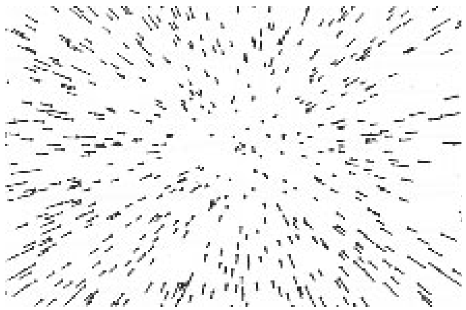
$$f = -\int_{S^2} \sigma_t(\eta) \frac{1}{\lambda(\eta)^2}(I - \eta\eta^\top) R_t (I - \eta\eta^\top) V \, d\theta_\eta \wedge d\psi_\eta$$

The angular torque applied by the dissipation

$$\tau = -\int_{S^2} \sigma_a(\eta) \eta_\times R_a \eta_\times \Omega \, d\theta_\eta \wedge d\psi_\eta$$

Optic flow is an visual cue that can be used with great effect for the control of aerial robotic vehicles.

Optical divergence is proportional of vertical velocity over distance.



(a) Hovering flight above the textured terrain

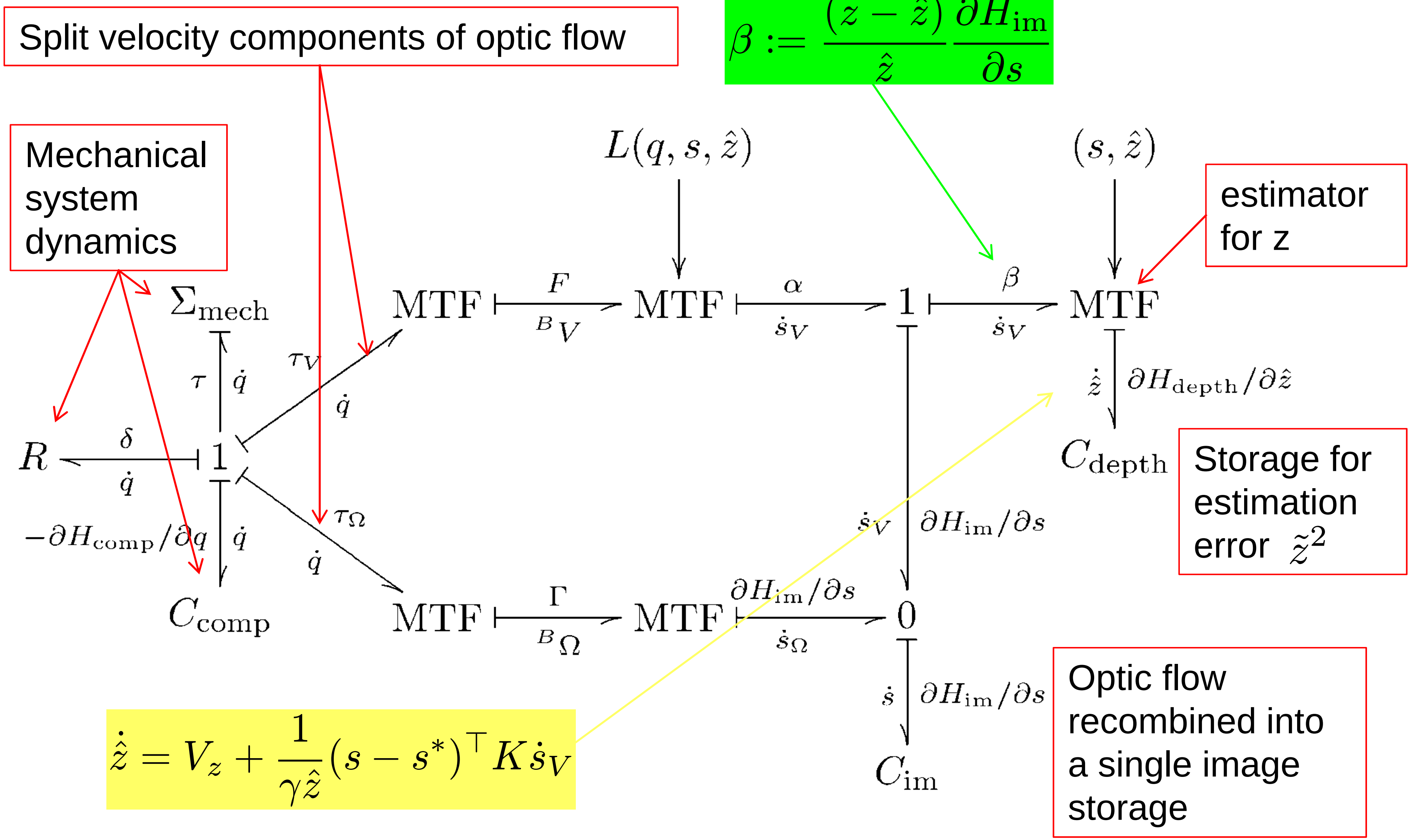
(b) Hovering flight around the textured corner

$$\tau_i = L_i^\top(s_i, z_i, q)e_i \quad \xmapsto[\dot{q}]{\tau_i} \text{MTF} \xmapsto[\dot{s}_i]{e_i} \quad \dot{s}_i = L_i(s_i, z_i, q)\dot{q}$$

- The Jacobian matrix $L_i(q, s_i, z_i)$ in an Image based visual servo scheme depends on the depth z_i of the image point with respect to the camera.
- The depth z_i are not directly measured due to the projective nature of imaging devices.
- The simplest work around to this issue is to use approximations $(\hat{z}_1, \ldots, \hat{z}_n)$ to the depth and compute an estimate $\hat{L} := L(q, s, \hat{z}_1, \ldots, \hat{z}_n)$ of the true Jacobian matrix.
- This has been found to be very effective in practice and IBVS is known to be highly robust to false estimation in depth.
- Nevertheless,it is of interest to develop a more detailed analysis using the port Hamiltonian framework.

Bond graph for dynIBVS with depth estimation

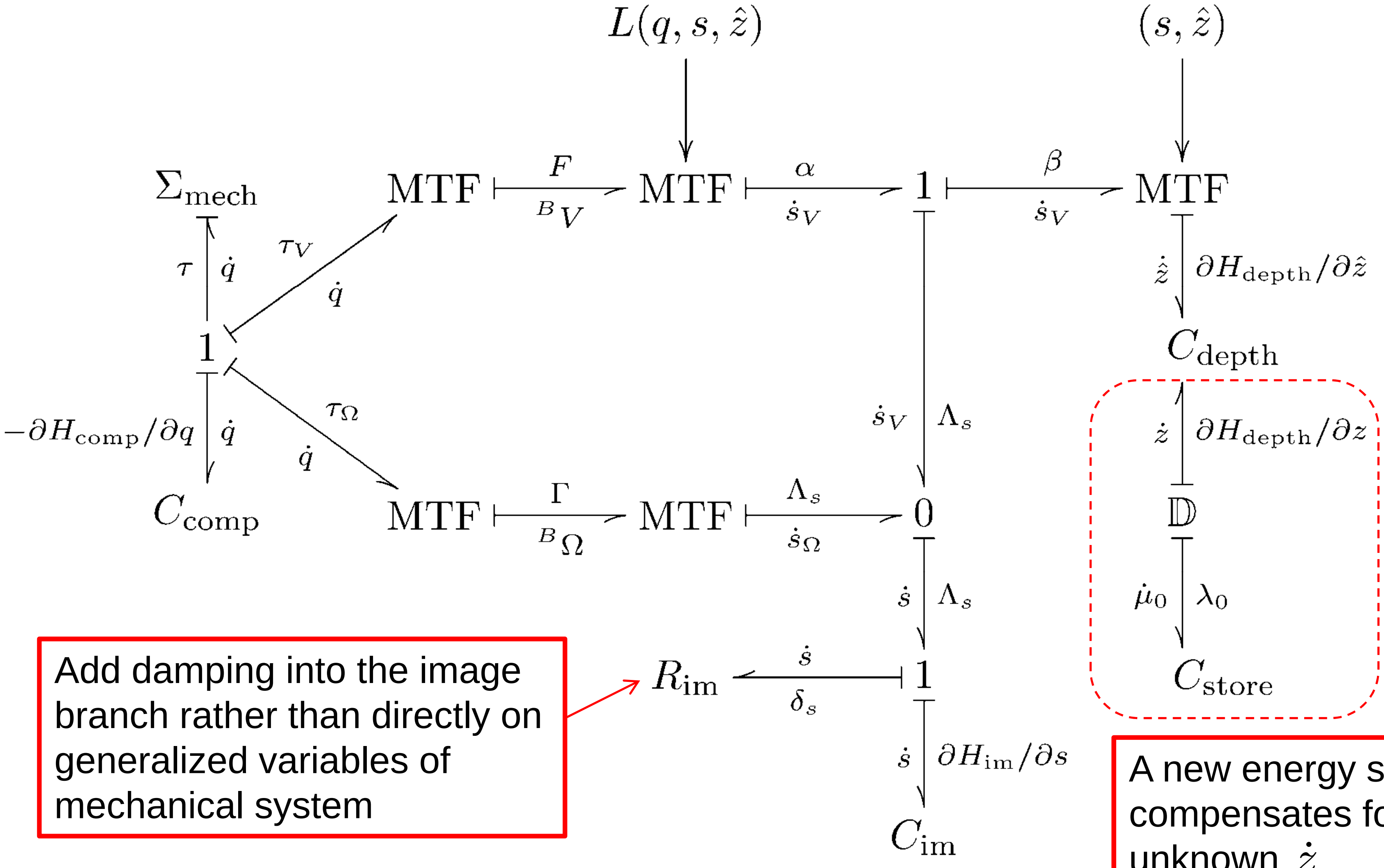
$$\beta := \frac{(z-\hat{z})}{\hat{z}}\frac{\partial H_{\mathrm{im}}}{\partial s}$$



$$\dot{\hat{z}} = V_z + \frac{1}{\gamma\hat{z}}(s-s^*)^\top K \dot{s}_V$$

dynIBVS with depth and velocity estimation

Australian National University



Add damping into the image branch rather than directly on generalized variables of mechanical system

A new energy store compensates for unknown $\dot{z}$

- Vision provides a rich and practical sensor modality for control of aerial vehicles
- To obtain high performance control of aerial vehicles, it is necessary to consider the dynamics of the vehicle.
- Existing results are all preliminary, particularly in the sense that existing lightweight, low-power vision processing technology is not well developed.

There are many opportunities for significant contribution remaining in this field.

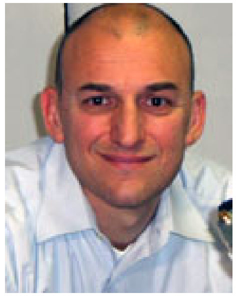
Collaboration


Tarek Hamel


Felix Schill


Nick Barnes

Odile Bourquardez



Stefano Stramigioli


Francois Chaumette


Nicolas Guenard


Jonathon Oh, Florent Le Bras, Eric Hou


Bruno Herisse


Chris McCarthy


Luke Cole


Francois-Xavier Russotto


Leigh Moffit


Peter Corke