4th IAA Conference on University Satellite Missions and CubeSat Workshop, Rome, Italy, 4-7 December 2017

Microsatellite Mock-up Control Using Reinforcement Learning Technique

Mark Shachkov, Danil Ivanov,
Mikhail Ovchinnikov
Keldysh Institute of Applied Mathematics, Moscow, Russia

ОРДЕНА ЛЕНИНА
ИМ.М.В.КЕЛДЫША
РАН

Machine Learning Techniques Application to Formation Flying Control

- **The machine learning improves the performance of the control algorithm**
  - in the case of significant unknown disturbances
  - changing parameters of the environmental forces or control actuators
- **The reinforcement machine learning**
  - provides online tracking of the changing parameters
  - requires the time to be trained

SPHERES mock-ups on board the ISS

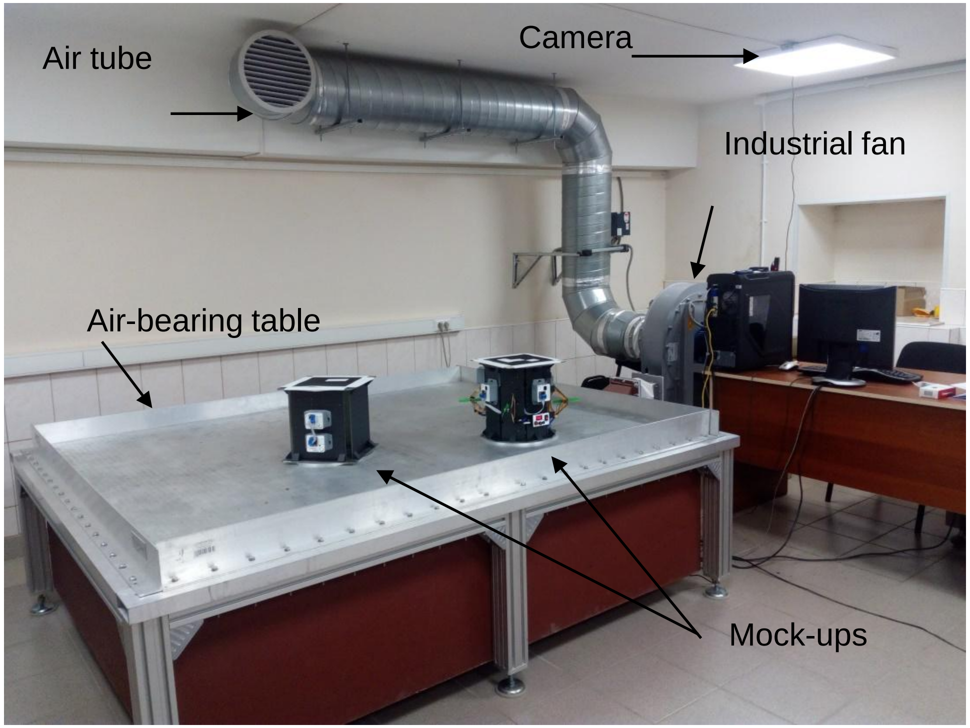
- **Examples of applications:**
  - nonlinear controller for deep-space spacecraft formation flying
  - improve the performance of the sliding mode control applied to the formation flying
  - MIT SPHERES control algorithm

Problem Statement

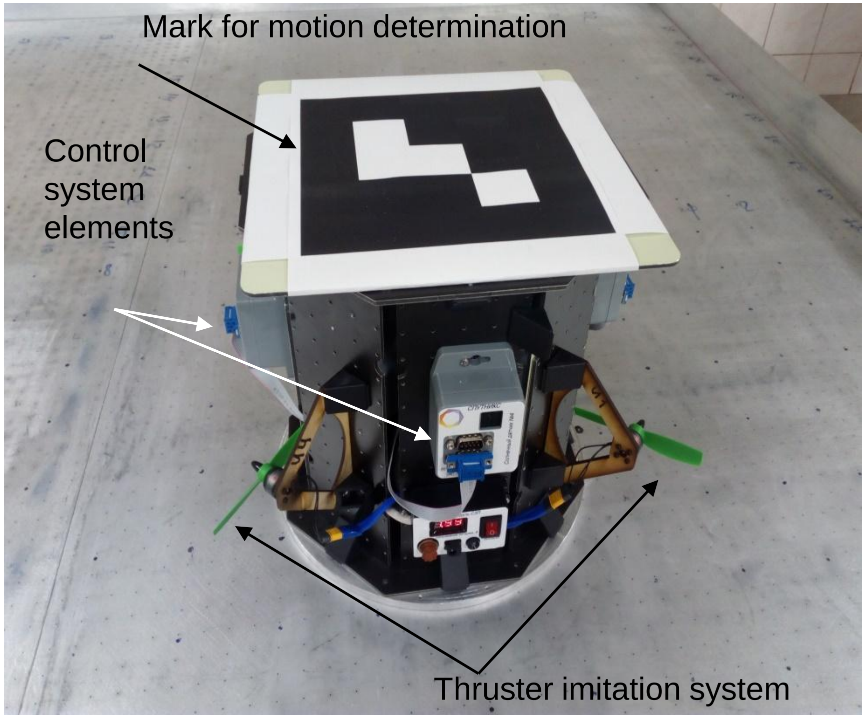
- Consider two microsatellites flying along the circular relative reference trajectory
- Unknown or ill-defined disturbances are unaccounted in the control algorithm
- The tracking error of the reference trajectory appears

- It is necessary to develop an adaptive control algorithm using reinforcement learning technique and test it using the microsatellite mock-ups on the air bearing test bench

Planar Air Bearing Test-Bench



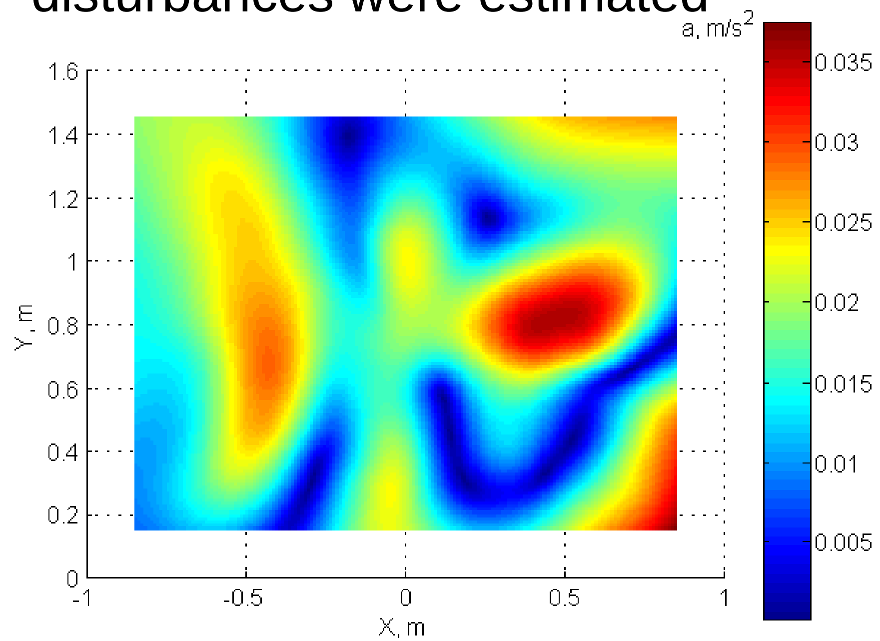
Test Bench COSMOS
(COmplex for Satellites MOtion Simulation)



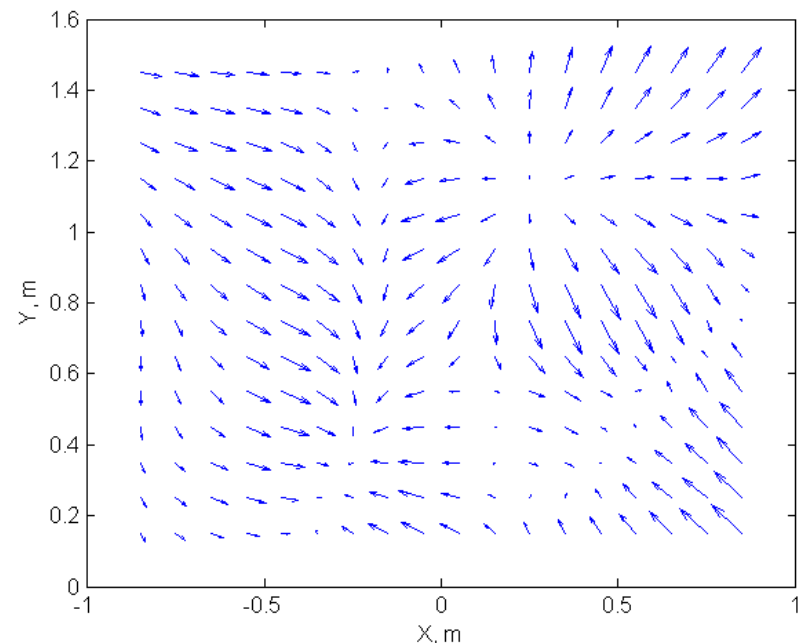
Microsatellite Mock-up

Preliminary Disturbances Determination on the Air Table

- Due to the uneven surface and non-uniform air flow along the table surface the disturbances appear
- Using set of the experiments of the free mock-up motion the disturbances were estimated



The magnitude of linear acceleration



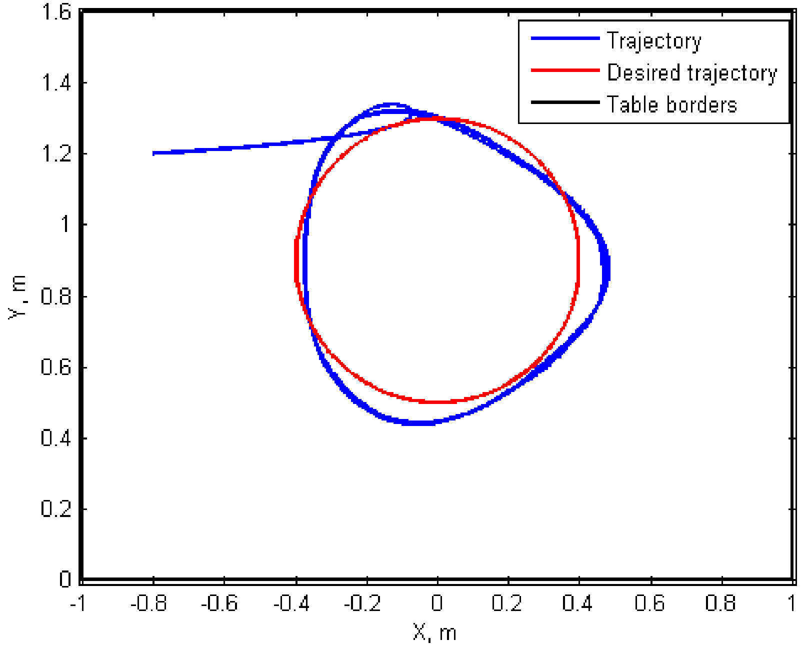
The direction of linear acceleration

The Mock-ups Controlled Motion

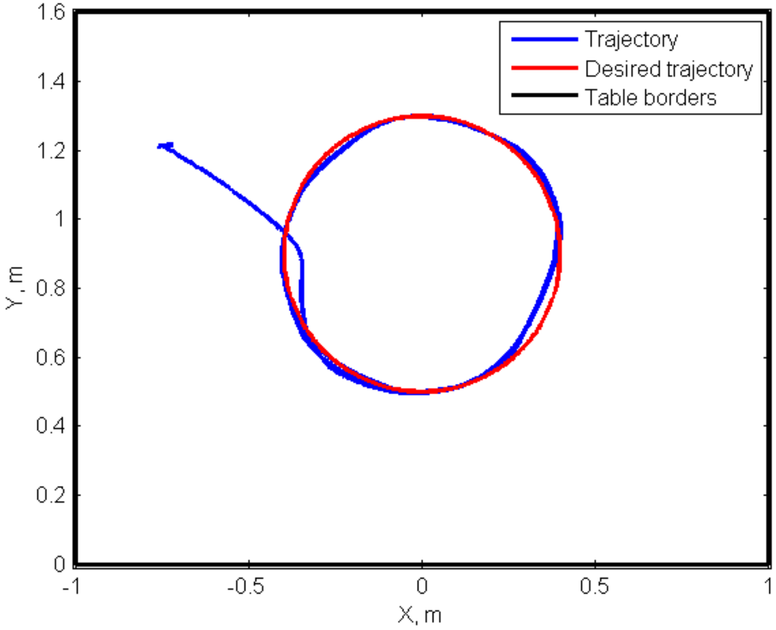
- The control algorithm for reference trajectory tracking

$$\mathbf{u} = -K_1\mathbf{e}_r - K_2\mathbf{e}_v + \ddot{\mathbf{q}}_d - \mathbf{d}$$

- The reference trajectory is circular



Motion under the control without the disturbances taken into account



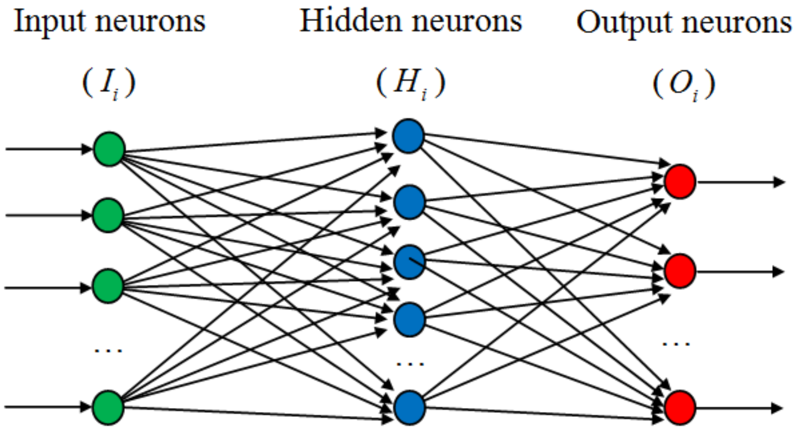
Motion under the control with the disturbances taken into account

Application of the Neural Network for the Disturbances Estimation

- The disturbances of the air table are known not accurately
- There are control realization errors
- Lets apply the neural network to estimate the disturbances in real time
- The three layer neural network is described by the equations

$$H_k = \sigma\left(b_k + \sum_{i=1}^{N_i} I_i w_{ki}\right),\ k = 1..N_h$$

$$O_k = \sigma\left(b_k^{'} + \sum_{i=1}^{N_h} H_i w_{ki}^{'}\right),\ k = 1..N_o$$



The three-layer perceptron neural network

- Consider the input and the output as

$$\mathbf{I} = \begin{bmatrix} \mathbf{q}^T & \dot{\mathbf{q}}^T & \mathbf{u}^T \end{bmatrix}^T \qquad \mathbf{O} = \mathbf{d}$$

The Reinforcement Learning

- Learning the neural network is setting the weights and biases

$$\xi = \left[w_{ki}, w_{ki}^{'}, b_k, b_k^{'} \right]^T$$

- The measurements of the position of the mock-up is the vector

$$\mathbf{z}(t) = [x \;\; y \;\; \varphi]^T$$

- Using the integration of the motion equations one can predict the measurements at the next step

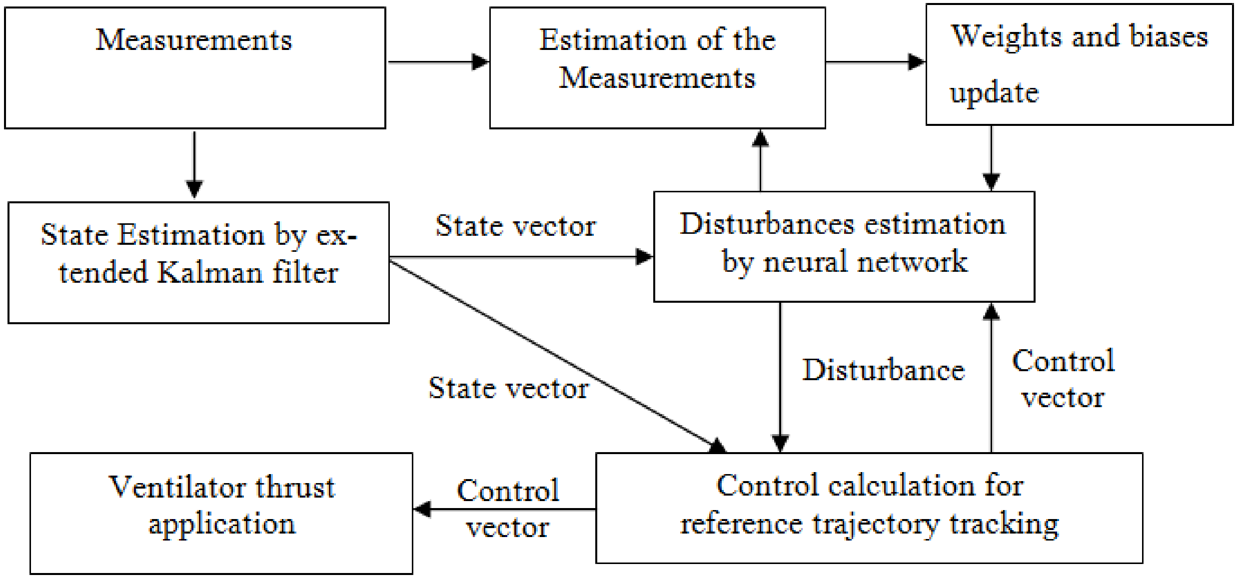
$$\tilde{\mathbf{z}}(t + \Delta t) = \int \mathbf{f}(\mathbf{q}) dt$$

- For neural network parameters update rule the value function is used

$$J(\xi) = \left(\mathbf{z}(t) - \hat{\mathbf{z}}(t) \right)^T \left(\mathbf{z}(t) - \hat{\mathbf{z}}(t) \right)$$

- The update rule is based on the "critic-only" method

Control System Loop



Dynamical characteristics of the mock-up

Mass of the whole mock-up,	5.2 kg
Mass of the flexible boom,	0.3 kg
Flexible boom length,	1.2 m
Mock-up body moment of inertia,	0.05 $kg*m^2$
Mock-up with booms moment of inertia,	0.15 $kg*m^2$
Natural main frequency,	1.5 Hz
Boom displacement vector,	[0.001;0.423]

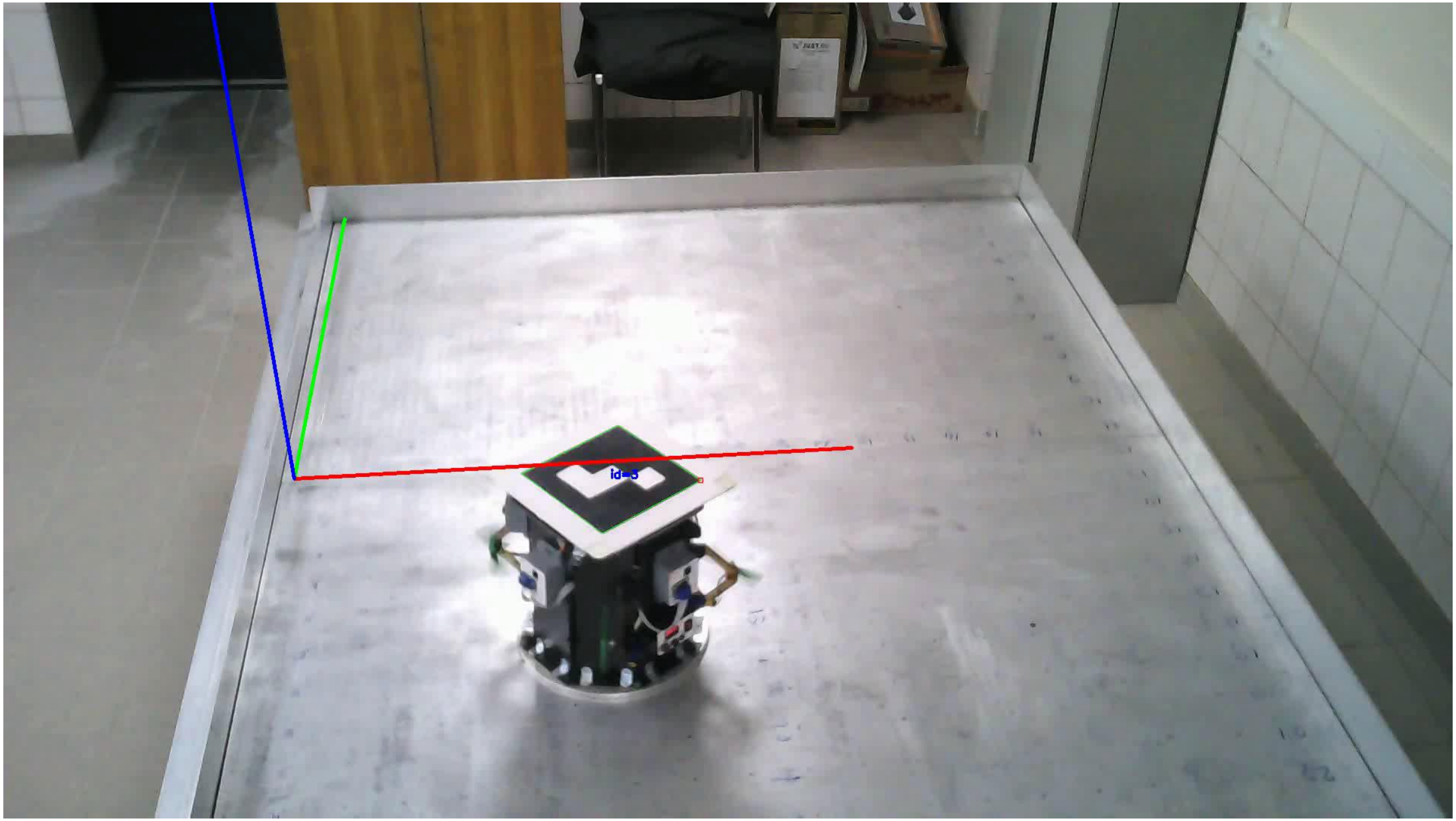
Control system parameters

Maximum ventilator thrust,	0.95 N
Maximum control force,	1.9 N
Maximum control torque,	0.4 N*m

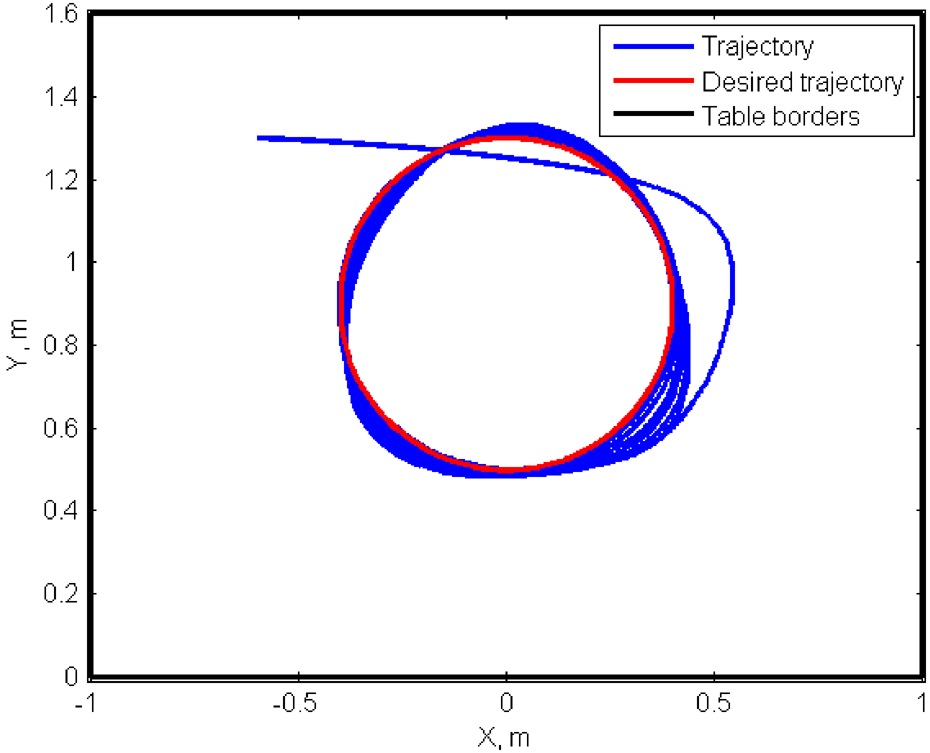
Measurement system parameters

Mean square position measurements error,	2 mm
Mean square angle measurements error,	0.1 deg

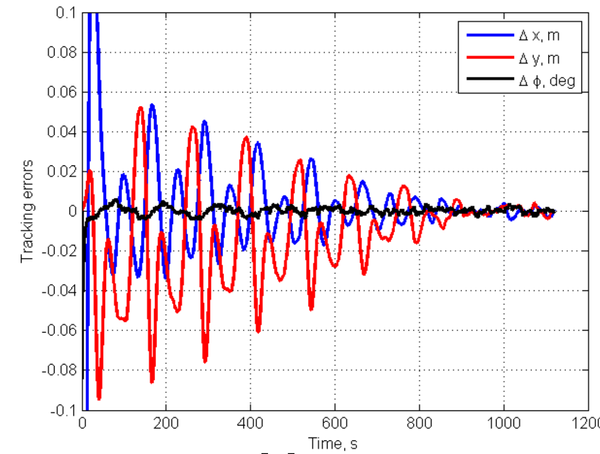
The Experiment



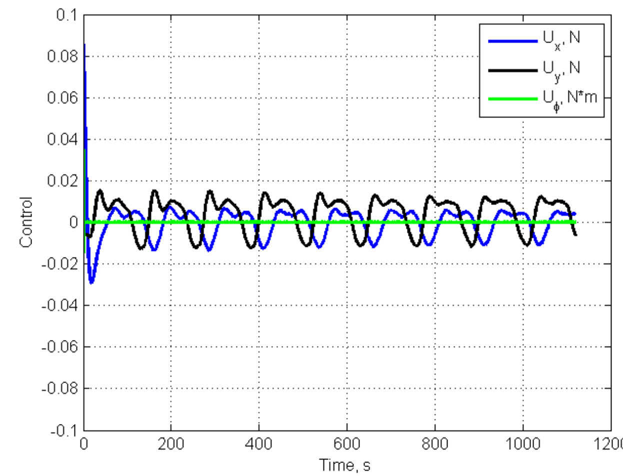
The Experimental Results



The mock-up motion during the machine learning process



Tracking errors



Calculated control

Conclusions

- The neural network real-time estimation of the disturbances acting on the microsatellites mock-up on the air bearing test-bench allows significant improvement of the performance of the tracking control algorithm
- The advantage of the reinforcement learning is that the developers are could not know accurately both the models of the mock-up motions and the test-bench disturbances, and nevertheless the controlled motion errors will be acceptable
- The disadvantage is that the neural network takes time to be trained and requires a computational power onboard for the real-time learning

Thank you for your attention!

Our web-site:
http://keldysh.ru/microsatellites/eng

Acknowledgment

- The work was supported by the Russian Foundation for Basic Research grants No. 17-01-00449, 16-01-00739 and 16-01-00634.