A GPU-powered Computational Framework for Efficient 3D Model-based Vision

Nikolaos Kyriazis, Iason Oikonomidis, Antonis A. Argyros

Institute of Computer Science, FORTH, Greece AND Department of Computer Science, University of Crete, Greece

PROBLEM

Provide efficient implementations for hypothesize-and-test vision methods that incorporate intense rendering as means of simulation.

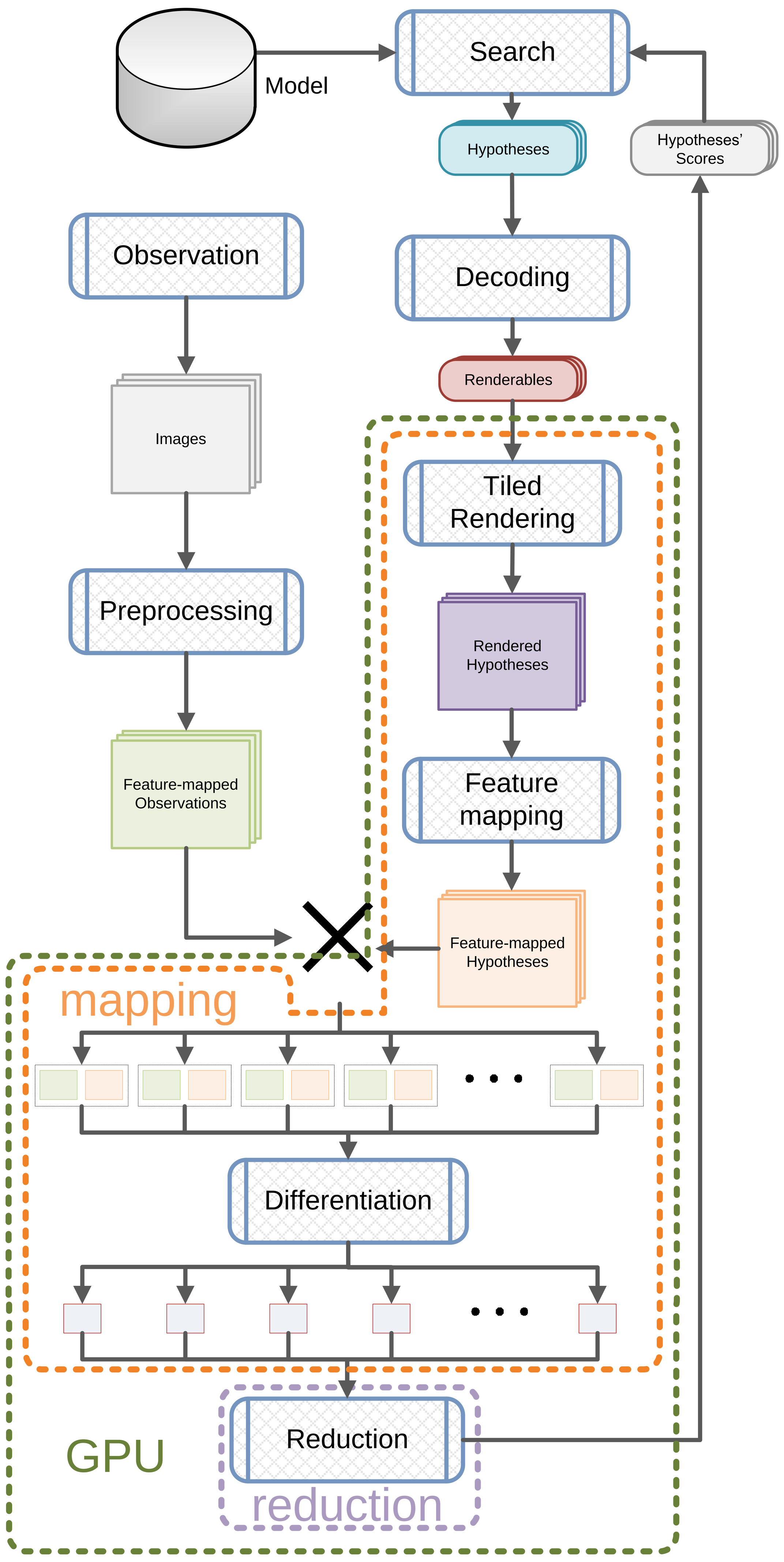
MOTIVATION

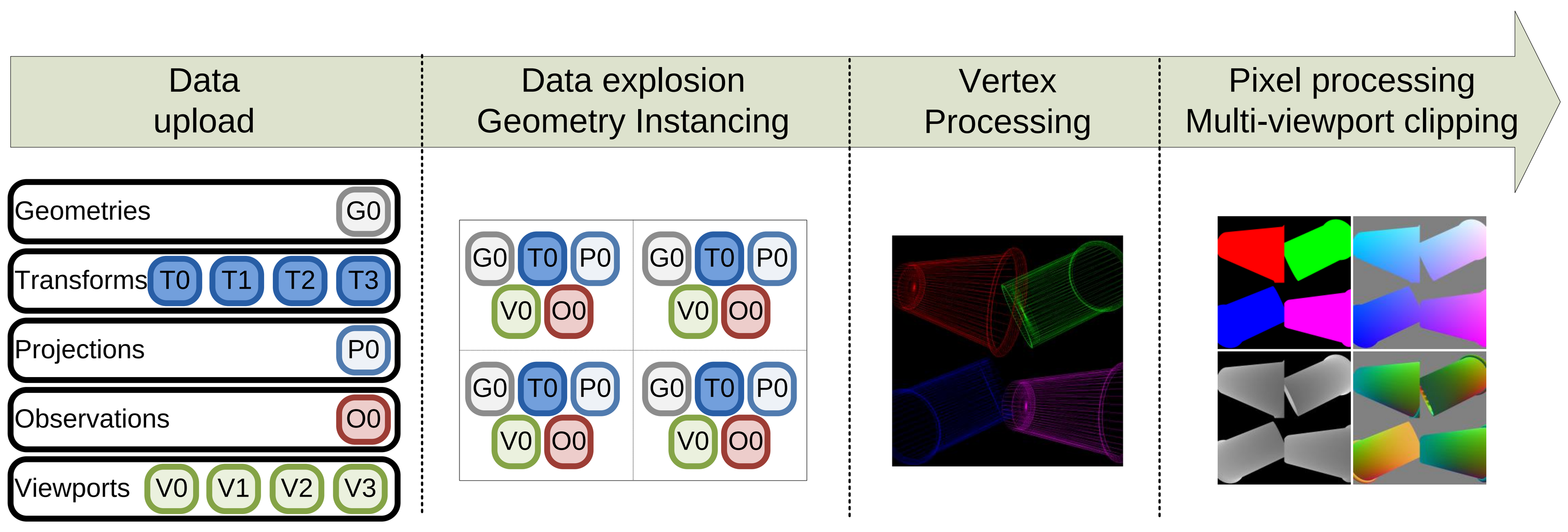
In computer vision, several problems are solved by employing hypothesize-and-test methods. Hypotheses can be made comparable to acquired images by means of 3D rendering.

MAIN IDEA

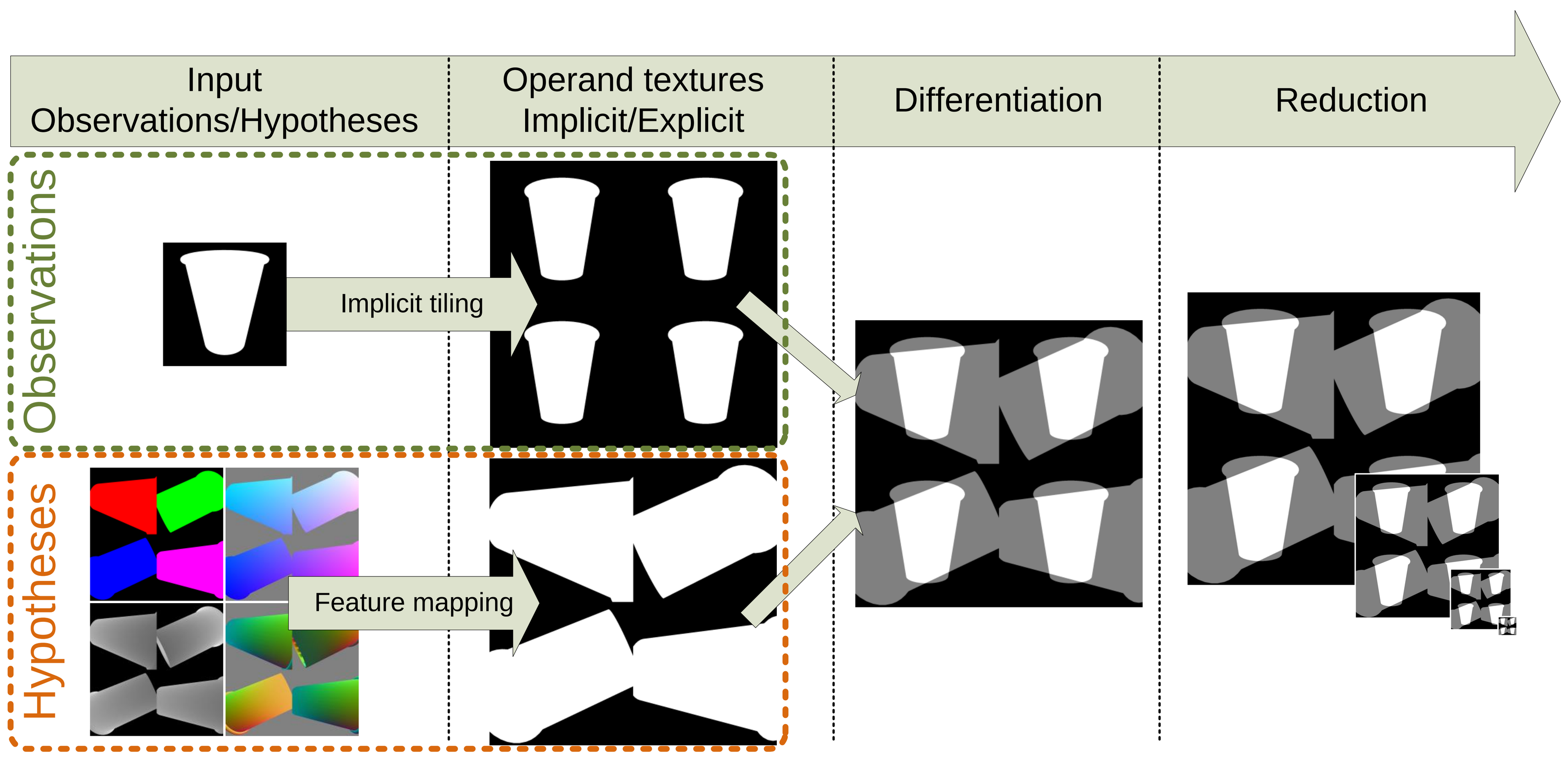
- 3D rendering is an inherently parallel process that is delegated to parallel hardware (GPUs)
- Parallel test/comparison criteria constitute the dominant case
- Exploitation of GPUs beyond traditional 3D rendering to satisfy the challenging computational demands of 3d model-based vision methods

METHOD





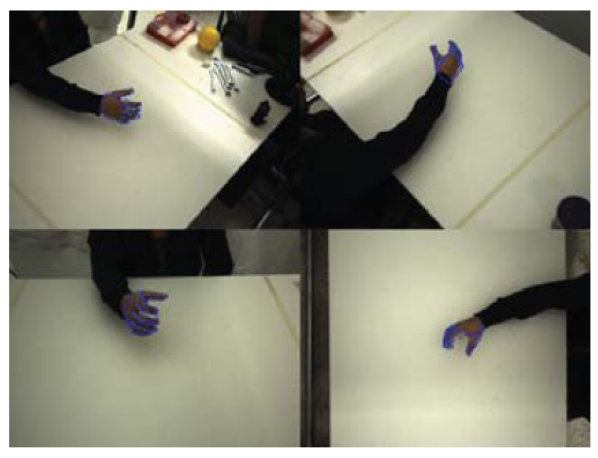
The tiled rendering process. Unique data are uploaded to the GPU, exploded into a tiled plan, processed in the vertex level and output in primary maps for later processing. Although there might be overlap of projected geometry across tiles during vertex processing this is remedied at the pixel-processing stage.



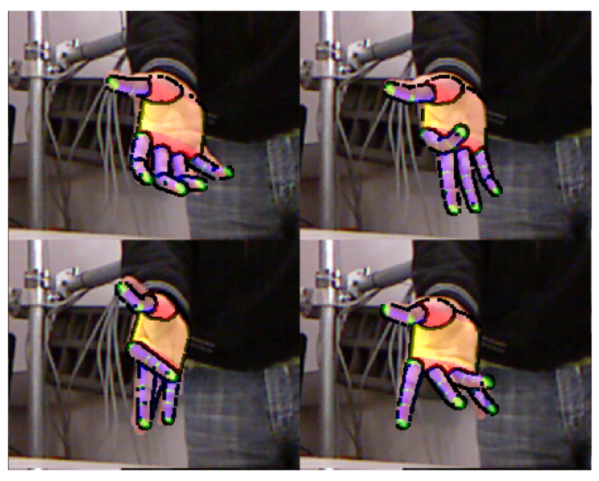
The differentiation process. Primary maps are mapped to the observations' feature space. Observations are implicitly tiled so as to match the tiled rendering of all hypotheses. A pixel wise differentiation is applied and the result is finally summed over the logical tiles by means of subsampling (data implosion).

APPLICATIONS

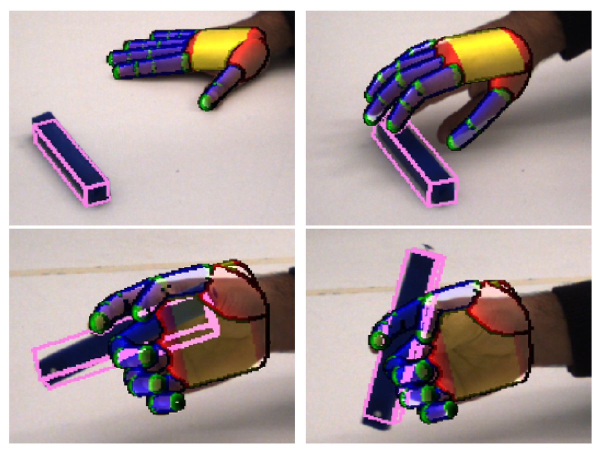
Tracking of "kinematic forests"



3D hand tracking from multiple cameras [1,2]
(2 fps for 4 cameras)

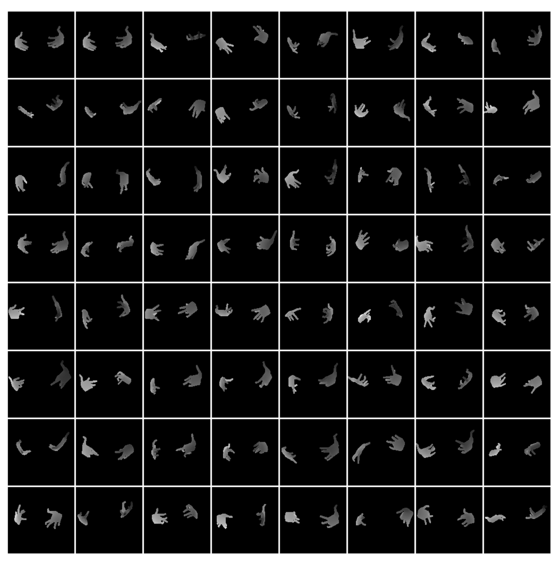


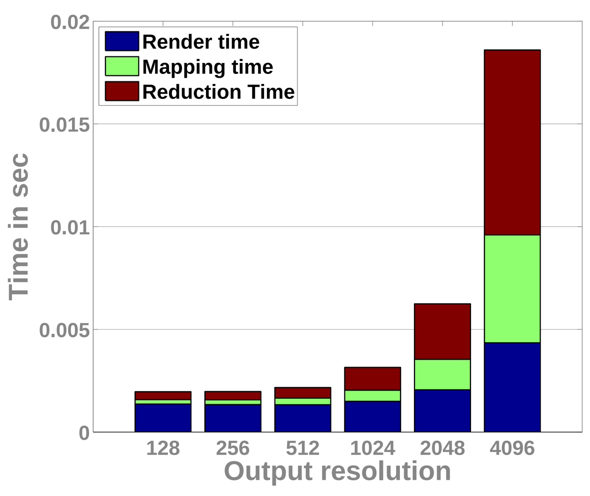
3D hand tracking from Kinect [1,3]
(15 fps for 1 sensor)

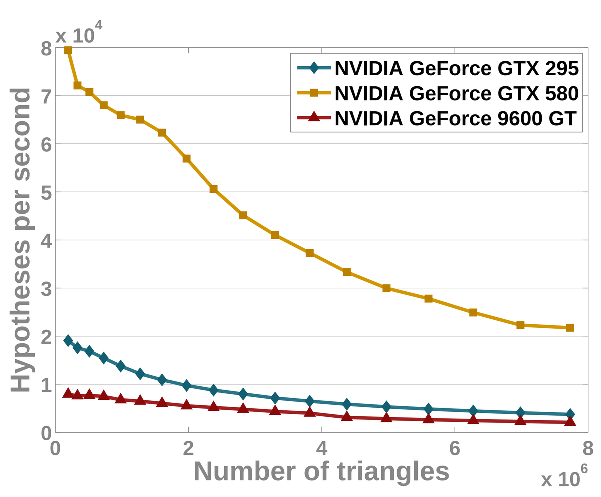


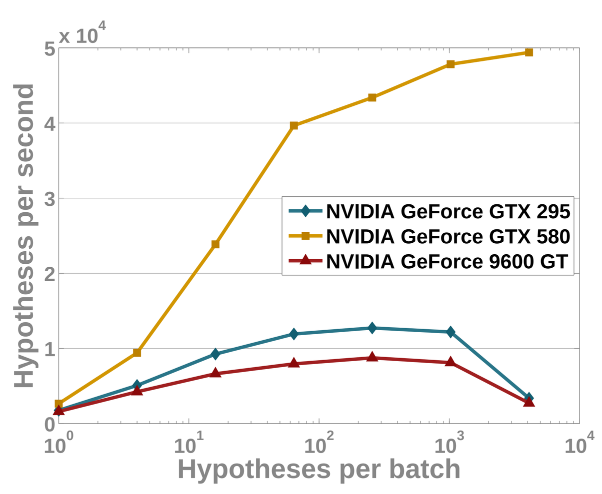
3D hand-object tracking from multiple cameras [1,4]
(2 fps for 4 cameras)

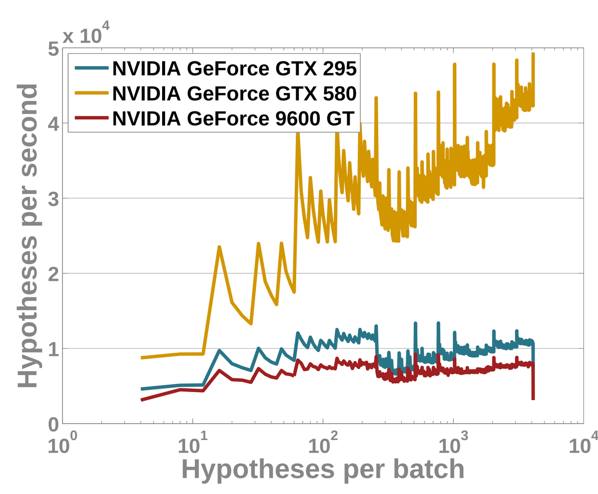
EXPERIMENTS











CONTRIBUTIONS

- Studied a challenging problem whose solution yields significant impact
- Identified a architecture with carefully designed modularity
- Presented an implementation that is based on GPU independent, commoditty pipeline, namely Direct3D 9
- Provided 3 distinct applications on the 3D articulated tracking problem

REFERENCES

[1] N. Kyriazis, I. Oikonomidis, and A. Argyros. A gpu-powered computational framework for efficient 3d model-based vision. Technical Report TR420, ICS-FORTH, July 2011.

[2] I. Oikonomidis, N. Kyriazis, and A. Argyros. Markerless and efficient 26-dof hand pose recovery. In *ACCV 2010*, pages 744–757. Springer, 2010.

[3] I. Oikonomidis, N. Kyriazis, and A. Argyros. Efficient model-based 3d tracking of hand articulations using kinect. In *BMVC 2011*. BMVA, 2011.

[4] I. Oikonomidis, N. Kyriazis, and A. Argyros. Full dof tracking of a hand interacting with an object by modeling occlusions and physical constraints. In *ICCV 2011*. IEEE, 2011.

MORE INFORMATION

Web: http://www.ics.forth.gr/~kyriazis/?e=10 E-mail: {kyriazis,oikonom,argyros}@ics.forth.gr

This work was partially supported by the IST-FP7-IP-215821 project GRASP