3D Object Class Recognition
Using the Web as Knowledge Repository

Walter Wohlkinger, Markus Vincze

[ww,vm] @acin.tuwien.ac.at

ACIN AUTOMATION & CONTROL INSTITUTE
INSTITUT FÜR AUTOMATISIERUNGS-
& REGELUNGSTECHNIK

TU WIEN TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology

robots @ home

Motivation

Object class recognition needed for robotics in home environment
Affordable robots with just a stereo camera as only sensor
Easy method required to teach robot new objects / object classes

web
learn ↕ share
interact
robot ⟷ user
teach
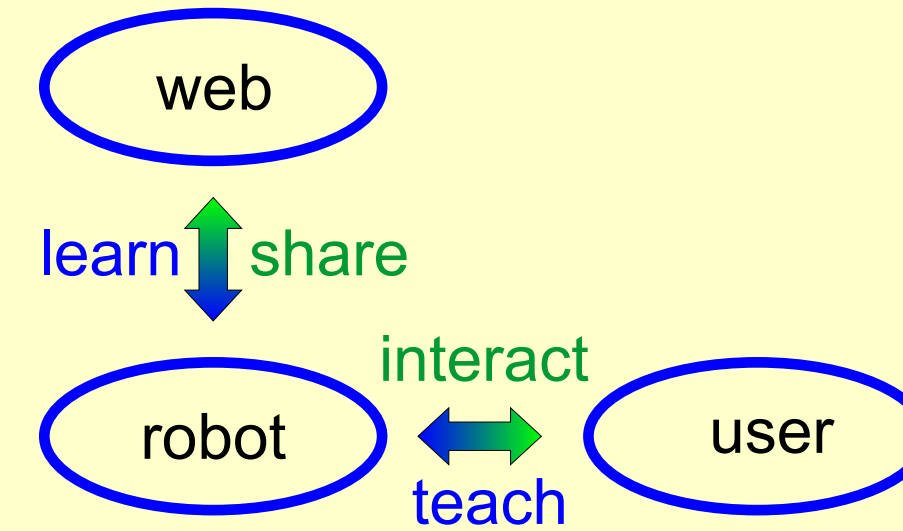
Objectives

User-friendly interaction/interface
Extensibility, scalability
Fast and unsupervised training
Handle large intraclass variability

Approach
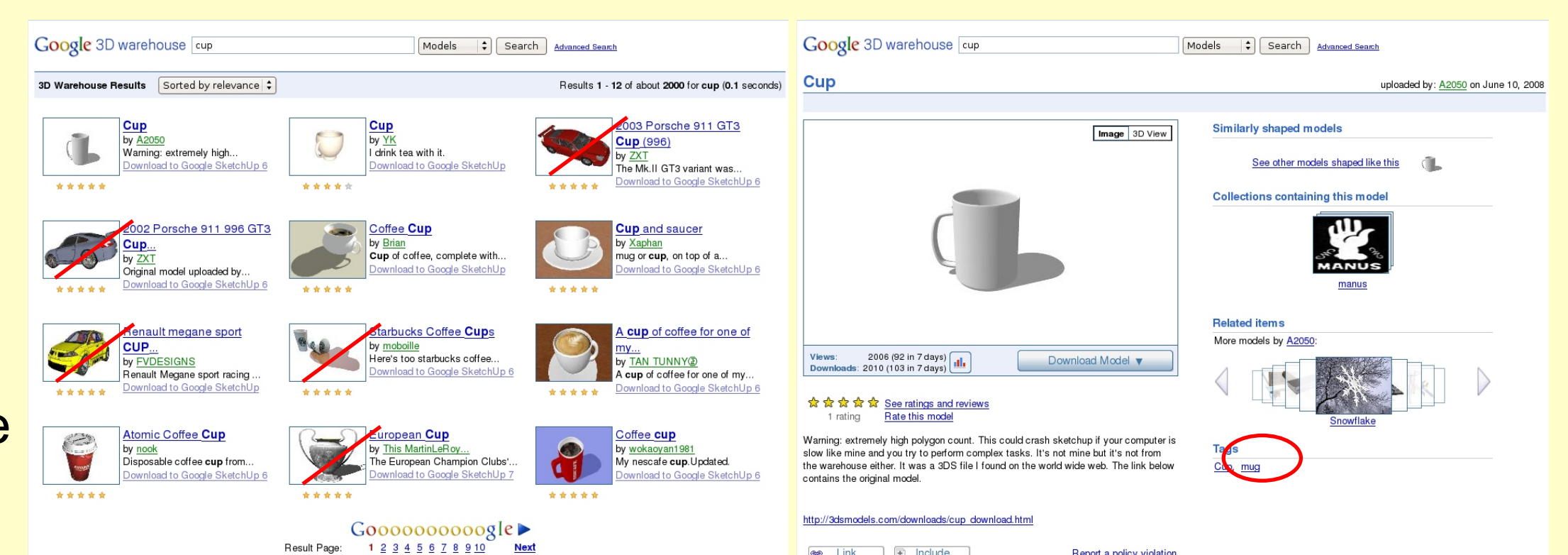
Model Aquisition: 3D model download from Google Warehouse by keyword, e.g. „dining chair"
Outlier Detection: discarding „wrong" models by comparing against images from Google Image Search and Walmart
Domain Adaption: generate 2.5D models from various views and sampling
Descriptor Generation: calculate 3D descriptor per view
**Matching**: nearest neighbour search, exemplar based classification

Challenges
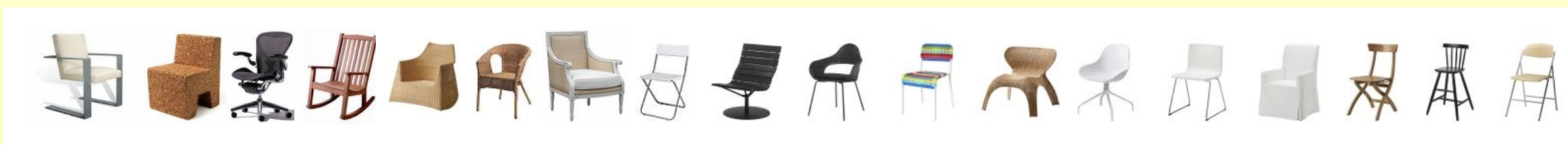
o Huge intra-class variability for some classes: large amount of training data required
o Incorrect labeled 3D models: vision algorithms required for filtering inappropriate models
o 3D models with extra geometry: filtering required, e.g. model of cup comes with a table in the scene
o View dependent descriptors increase model count: fast matching needed

Model Aquisition & Outlier Detection

example keyword: „dining chair„

o Make use of already prefiltered, labeled, up-to-date and free sources
o 3D models from Google Warehouse, reference-images from Google and Walmart
o Render Images from multiple views
o Calculate IS-Match descriptor [Donoser, 2009]
o Remove outliers from 3D model database with dissimilarity > threshold
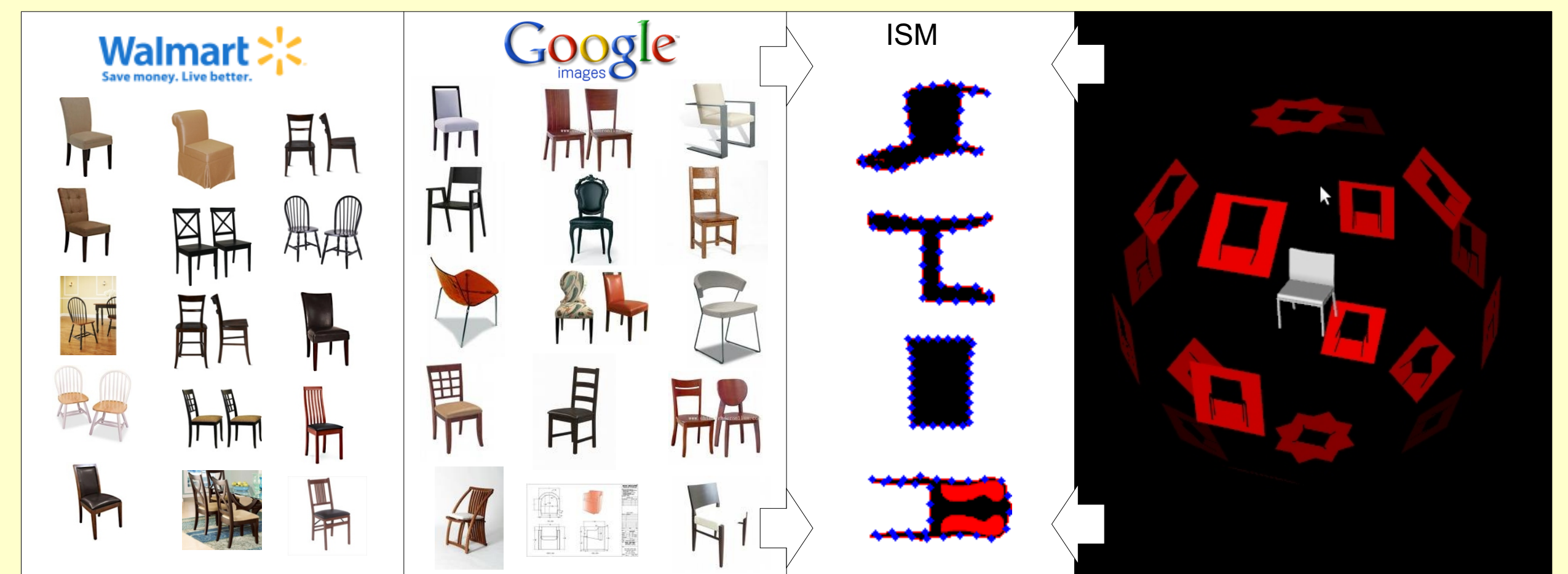
Domain Adaption and Descriptor Generation

o Adapt 3D models to dense stereo domain: generate 2.5D point clouds
o Views around model: 45° steps in elevation and azimuth
o Low resolution sampling for efficiency and generalization
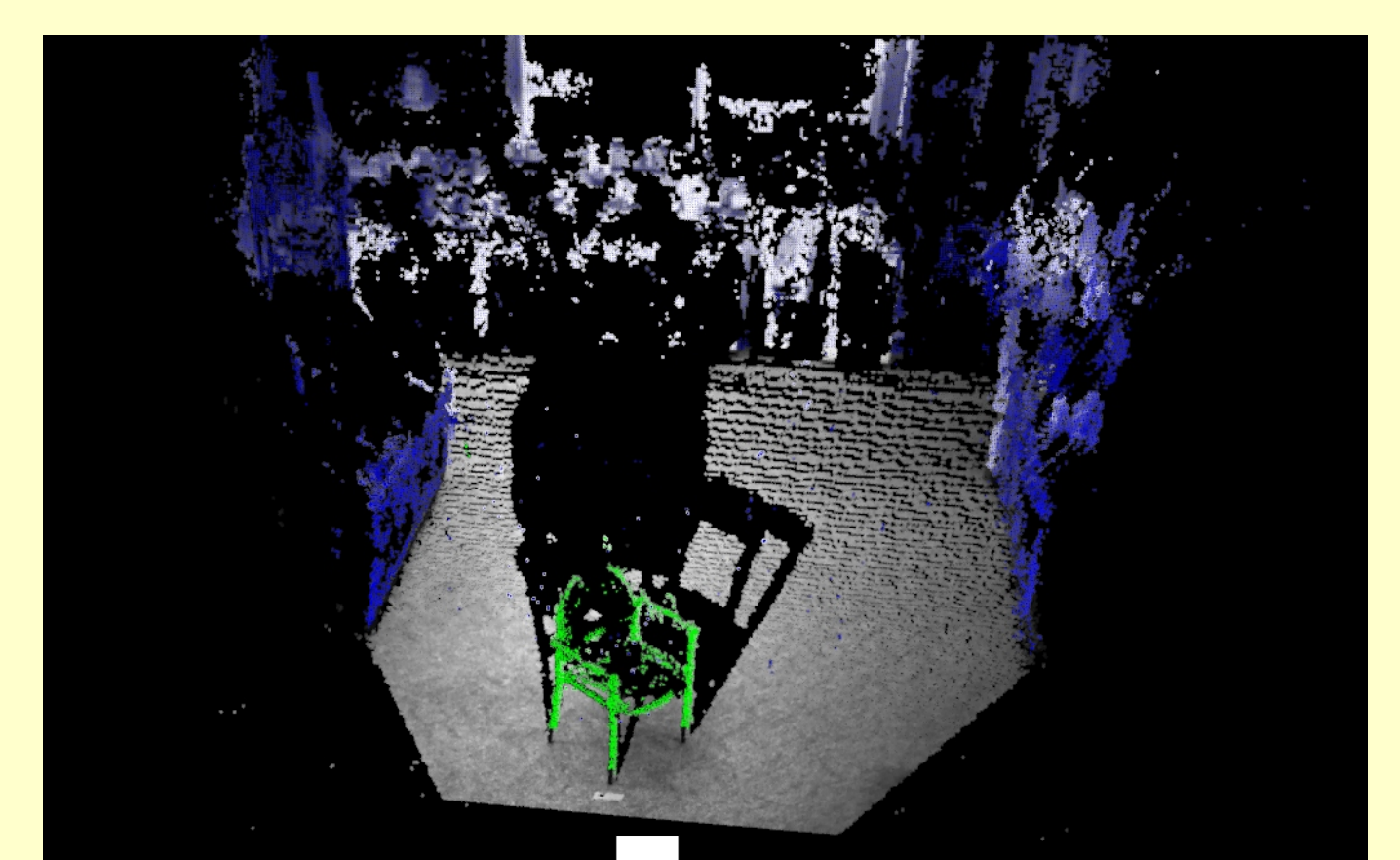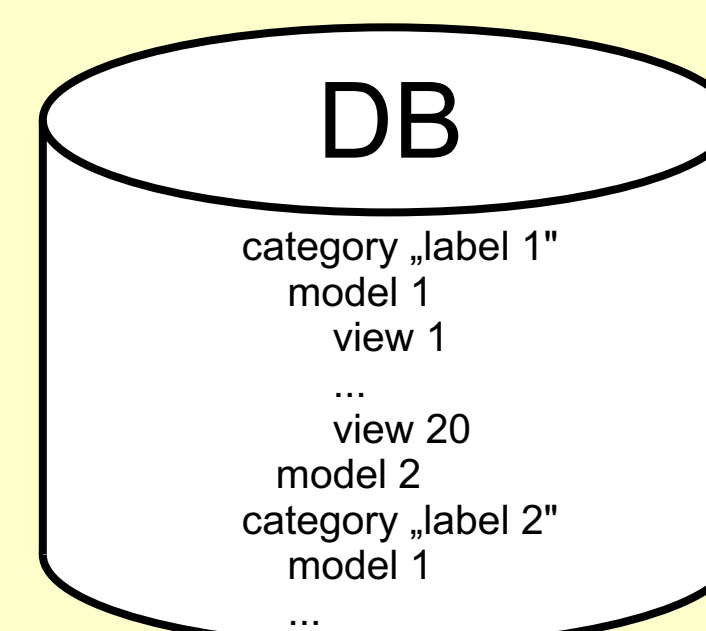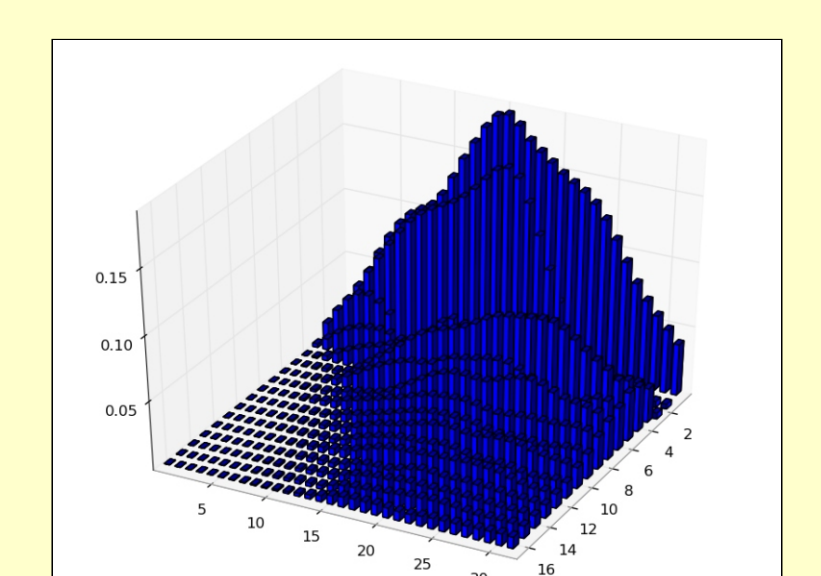o Spherical Harmonics Descriptor [Kazhdan, Funkhouser, 2002] for each view
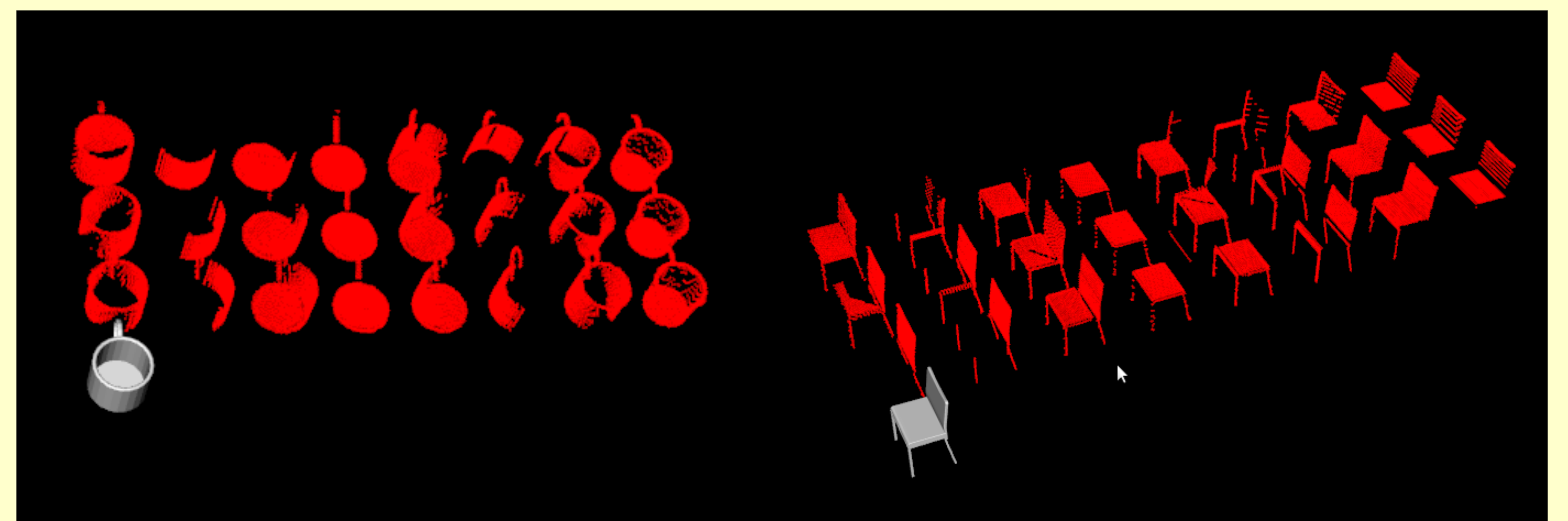
Matching

o Real time dense stereo algorithm [AIT]
o Filtering
o Support plane detection: prior for objects on ground plane, table, counter
o Segmentation and candidate extraction

for each candidate:
  o Calculate Spherical Harmonics Shape Descriptor
  o Compare against all models in Database: L2 norm of 2D histogram
  o Find best match(es) and return label of assigned category
  o Matching < 1sec

„class dining chair"

DB
category „label 1"
model 1
view 1
...
view 20
model 2
category „label 2"
model 1
...

size prior

Conclusion & Future Work

o Fast and modular framework for 3D object class recognition
o Easy to add categories: user inputs name of category by text or speech
o Fast to „learn" new categories

o Include IS-Match Shape descriptor in matching stage (to cope with sparse 3D stereo data)
o Increase detection performance through combination of descriptors

robots@home