

Robot Vision: A Holistic View

Contents

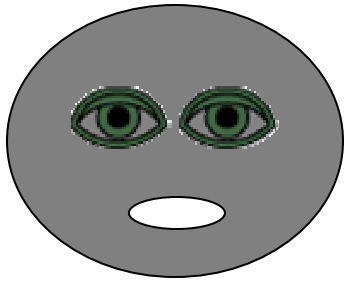
Introduction

3D-to-3D Matching

2D-to-3D Matching

Future Challenge

Model-based Vision



Contents

- Introduction

Introduction

- 3D-to-3D Matching

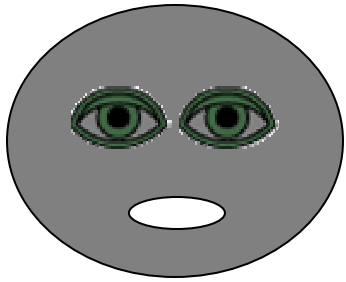
3D-to-3D Matching

- 2D-to-3D Matching

2D-to-3D Matching

Future Challenge

- Future Challenge



Contents

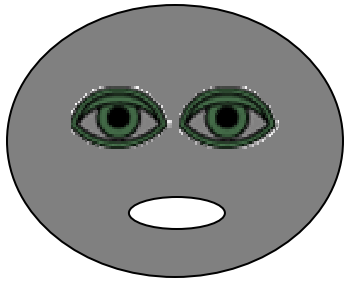
Introduction

- Model-based vision studies the aspect of inferring the occurrences of 3D objects from 2D images or 3D data. We distinguish it from (2D) pattern recognition.
- There are two motivations behind the development of model-based vision.
- The first motivation is object recognition, which is considered to be the next step, after the success of instrumental vision, or the (partial) success of reconstructive vision.
- The second motivation is image or photo interpretation, which human vision could easily cope with, in the absence of 3D data.

3D-to-3D Matching

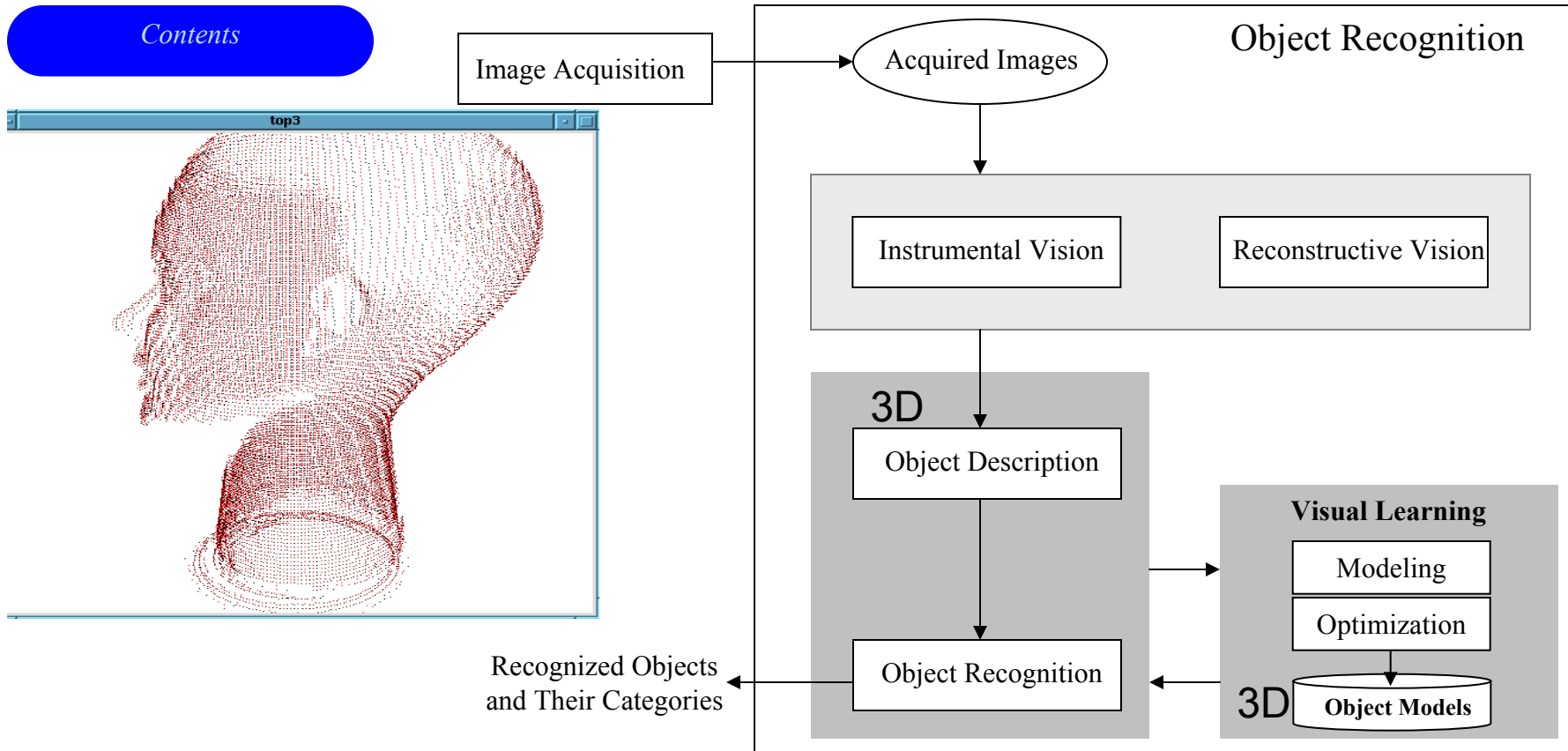
2D-to-3D Matching

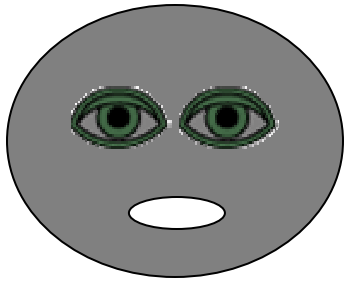
Future Challenge



Introduction

Motivation 1: Object Recognition as the Next Step after 3D Reconstruction

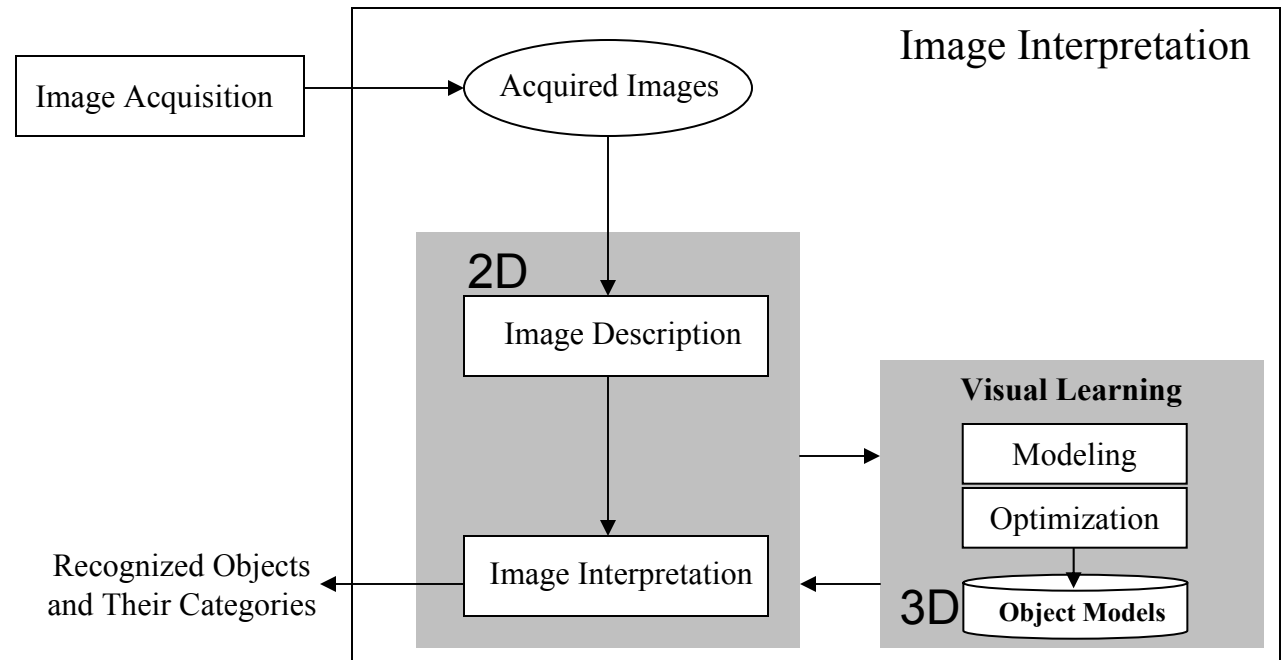
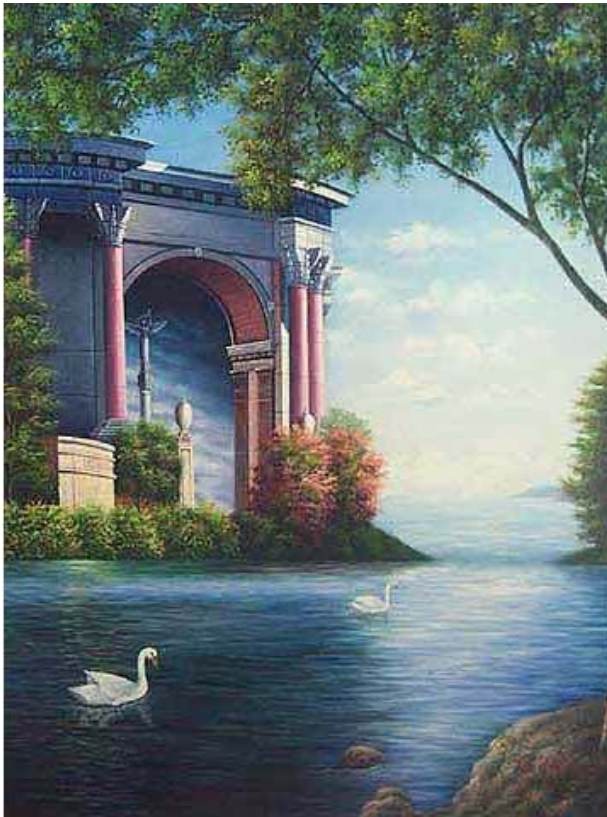


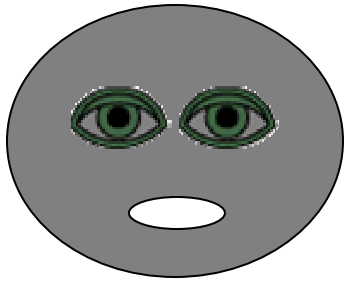


Introduction

Motivation 2: Image interpretation is a skill of human vision, which a robot should have.

Contents





Contents

Introduction

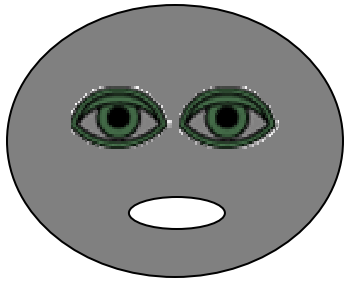
3D-to-3D Matching

2D-to-3D Matching

Future Challenge

Introduction

- For object recognition, 3D-to-3D matching methods have been investigated.
- For image interpretation, 2D-to-3D matching methods have also been investigated.
- However, there is no commonly accepted methodology in model-based vision.
- This explains the “stagnation” in computer, or robot, vision, despite very active research in this field.
- Therefore, there is an opportunity for any new ground, on which robot vision could achieve both image interpretation and object recognition within a same framework.



Contents

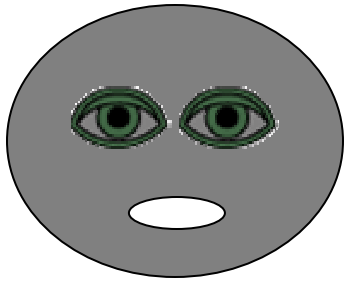
Introduction

2D-to-3D Matching

Future Challenge

3D-to-3D Matching

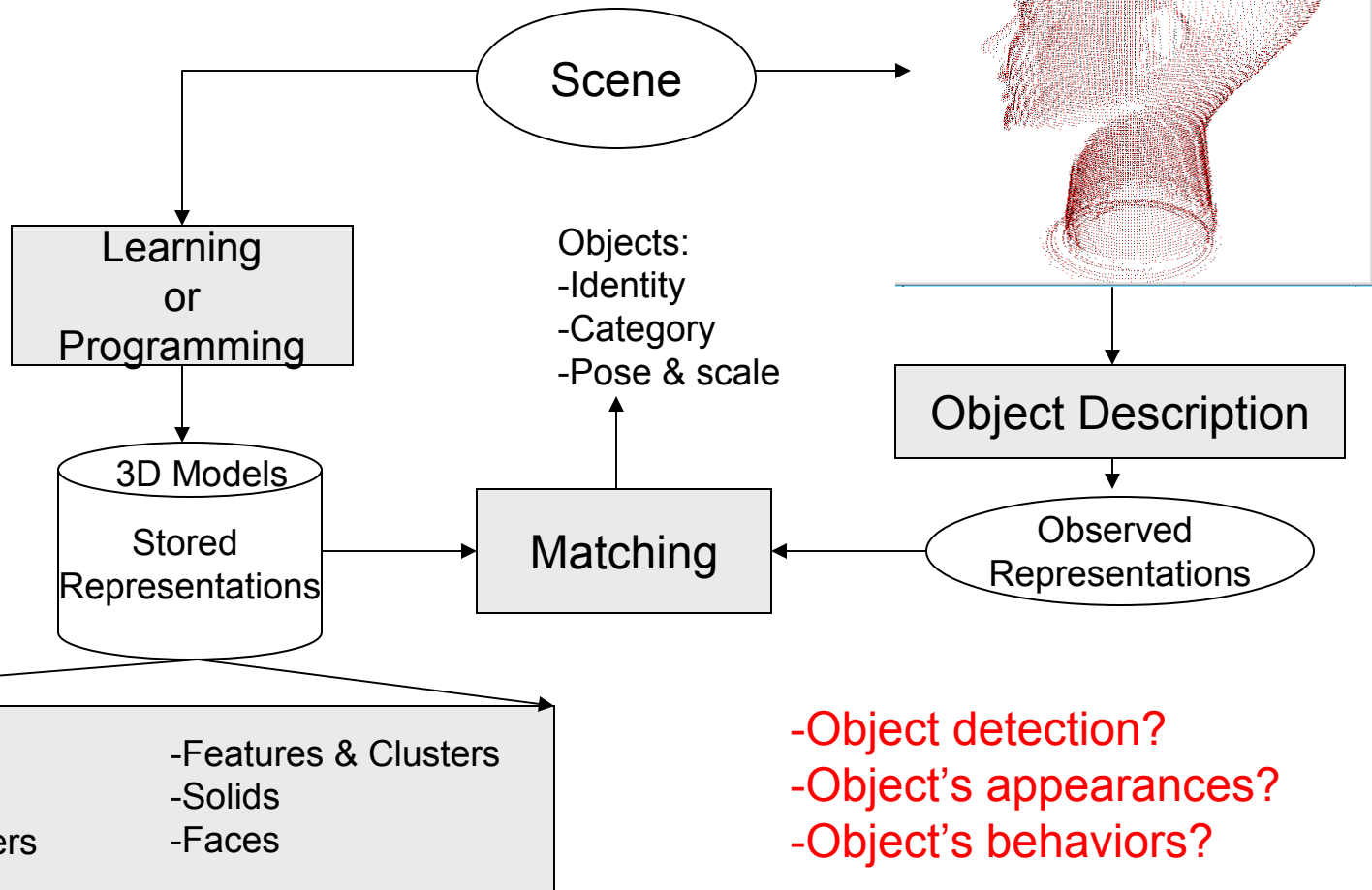
- 3D-to-3D matching methodology was pioneered by Nevatia & Binford (1977), followed by Oshima & Shirai (1983), Bolles et al (1983), Fisher (1983), Grimson & Lozano-perez (1984), and Faugeras & Hebert (1986), and many others.
- However, 3D-to-3D matching has very limited success because:
 - A) 3D reconstruction and object detection in real situation is still problematic.
 - B) Object modeling is still too simplistic.
 - C) There is no consideration of object “semantics” in a dynamic scene.
- These shortcomings appear obvious in the schematic diagram.

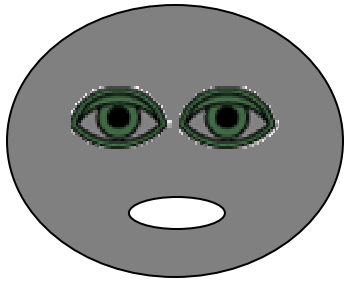


3D-to-3D Matching

Schematic Diagram

- Contents
- Introduction
- 3D-to-3D Matching
- 2D-to-3D Matching
- Future Challenge





Contents

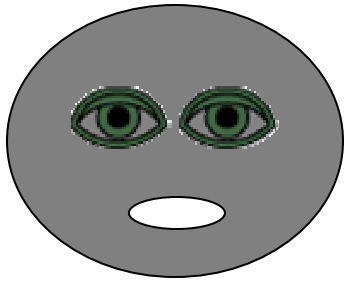
Introduction

3D-to-3D Matching

Future Challenge

2D-to-3D Matching

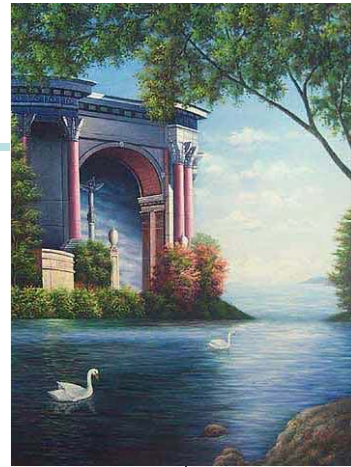
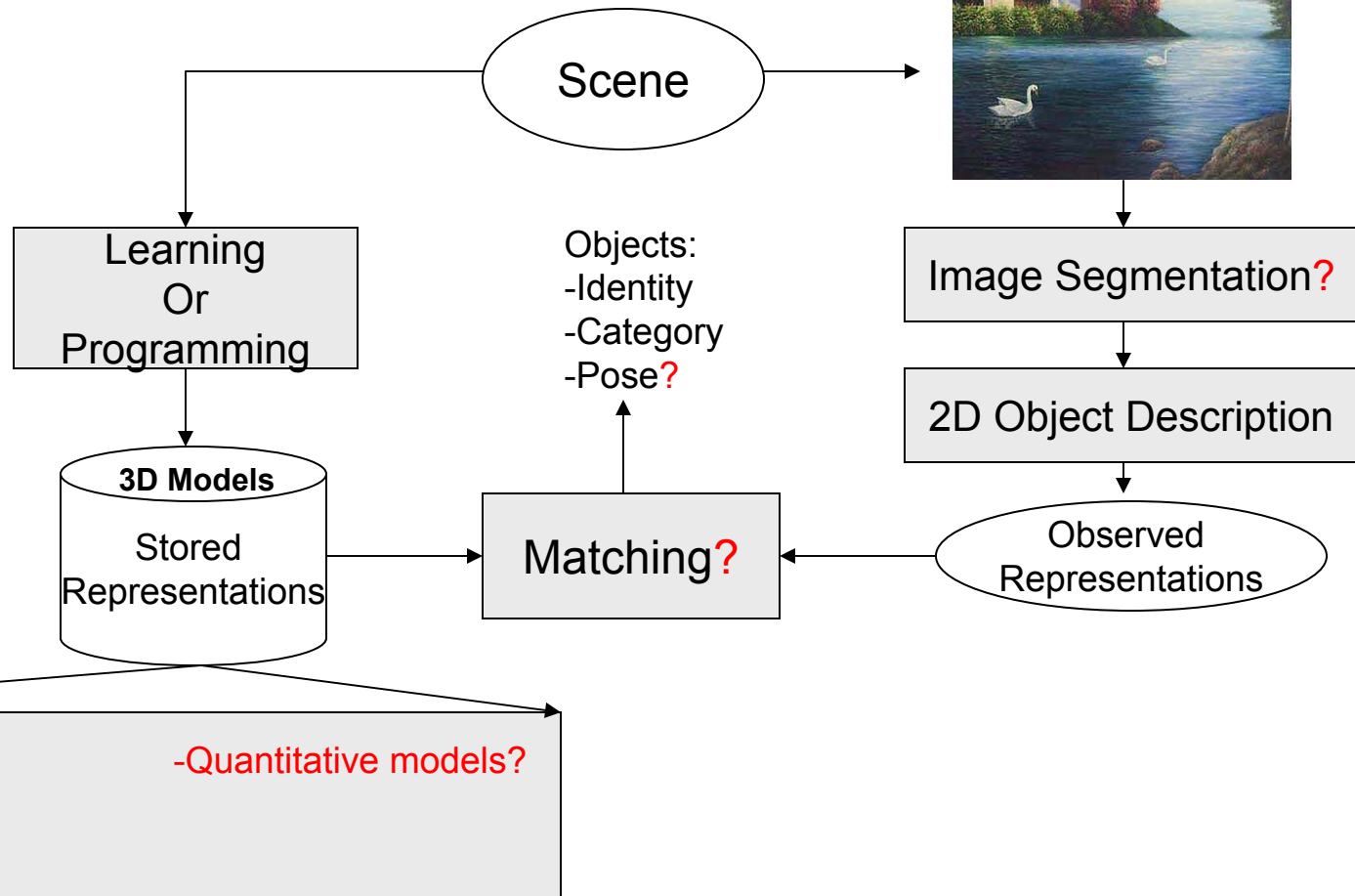
- 2D-to-3D matching methodology was pioneered by Brooks (1979) and Lowe (1985).
- Unfortunately, 2D-to-3D matching is not actively pursued because:
 - A) Image segmentation in real situation is still problematic.
 - B) It does not make much sense to match a 2D representation against a 3D representation (i.e. model).
 - C) The difficulty of inferring the “pose” of objects in 3D space.
- These shortcomings appear obvious in the schematic diagram.

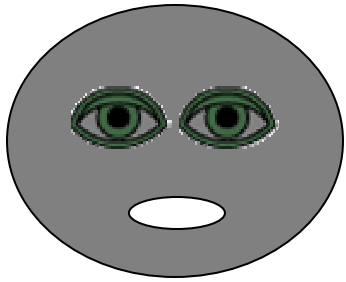


2D-to-3D Matching

Schematic Diagram

- Contents
- Introduction
- 3D-to-3D Matching
- 2D-to-3D Matching
- Future Challenge





Future Challenge

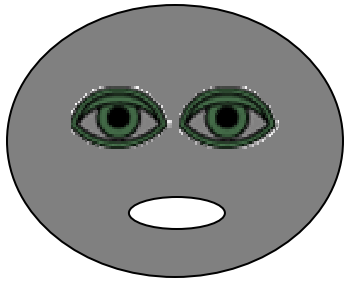
Contents

Introduction

3D-to-3D Matching

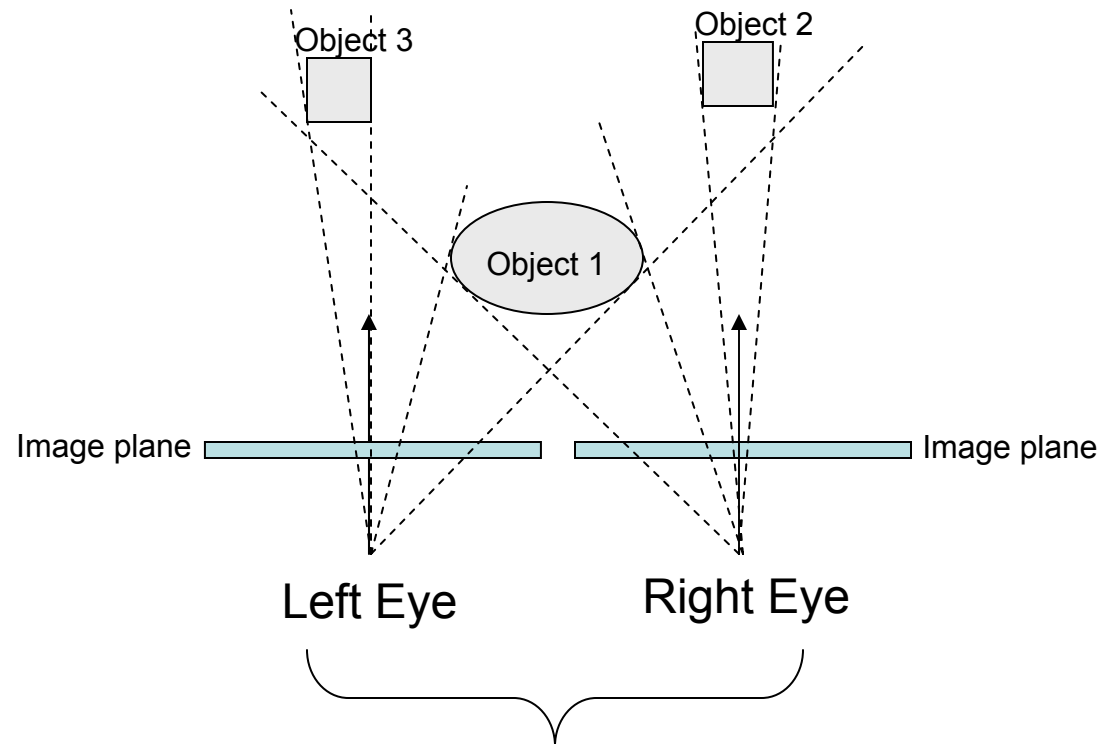
2D-to-3D Matching

- We know that 3D-to-3D matching is still problematic.
- We also know that 2D-to-3D matching is still difficult.
- Should we lose interest in model-based vision? Probably no! This is because:
 - Human vision is robust in the presence of occlusion, because each of our eyes could perform independently. This indicates that human vision is cooperative.
 - In addition, human vision is powerful enough to effortlessly interpret images, and recognize familiar objects. This suggests that vision should be model-based, or even knowledge-based (i.e. cognitive vision).

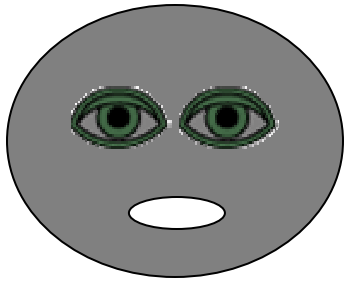


Future Challenge

Assumption 1: Human vision is robust in the presence of occlusion.

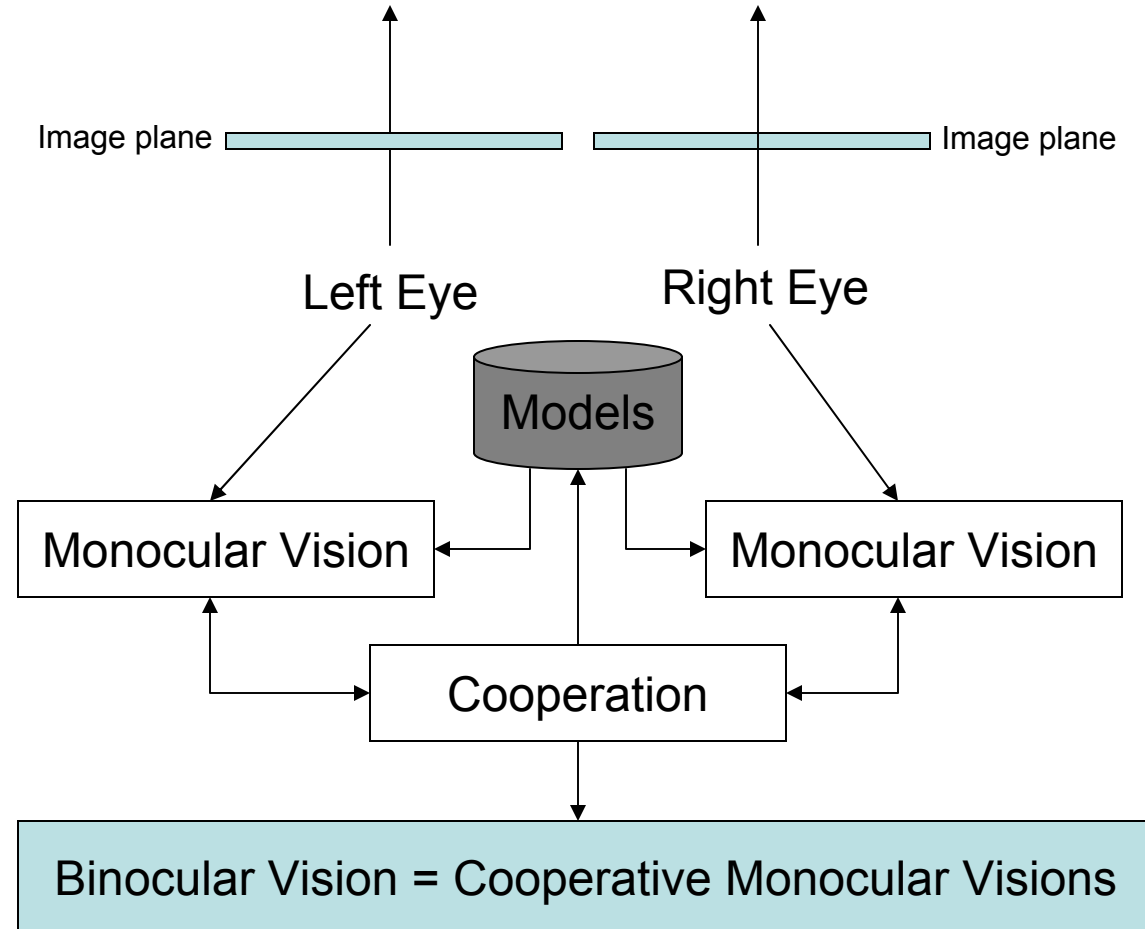


- Left eye only sees objects 1 and 2.
- Right eye only sees objects 1 and 3.
- The final result will still be three objects!



Future Challenge

Assumption 2: Human vision is cooperative, and is model-based.



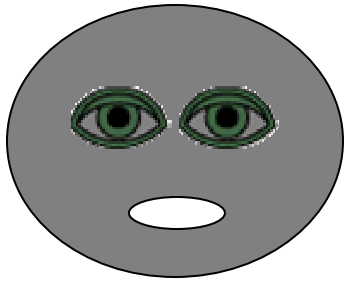
Contents

Introduction

3D-to-3D Matching

2D-to-3D Matching

Future Challenge



Contents

Introduction

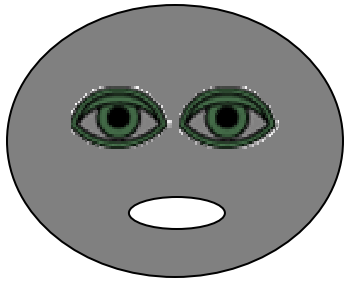
3D-to-3D Matching

2D-to-3D Matching

Future Challenge

Future Challenge

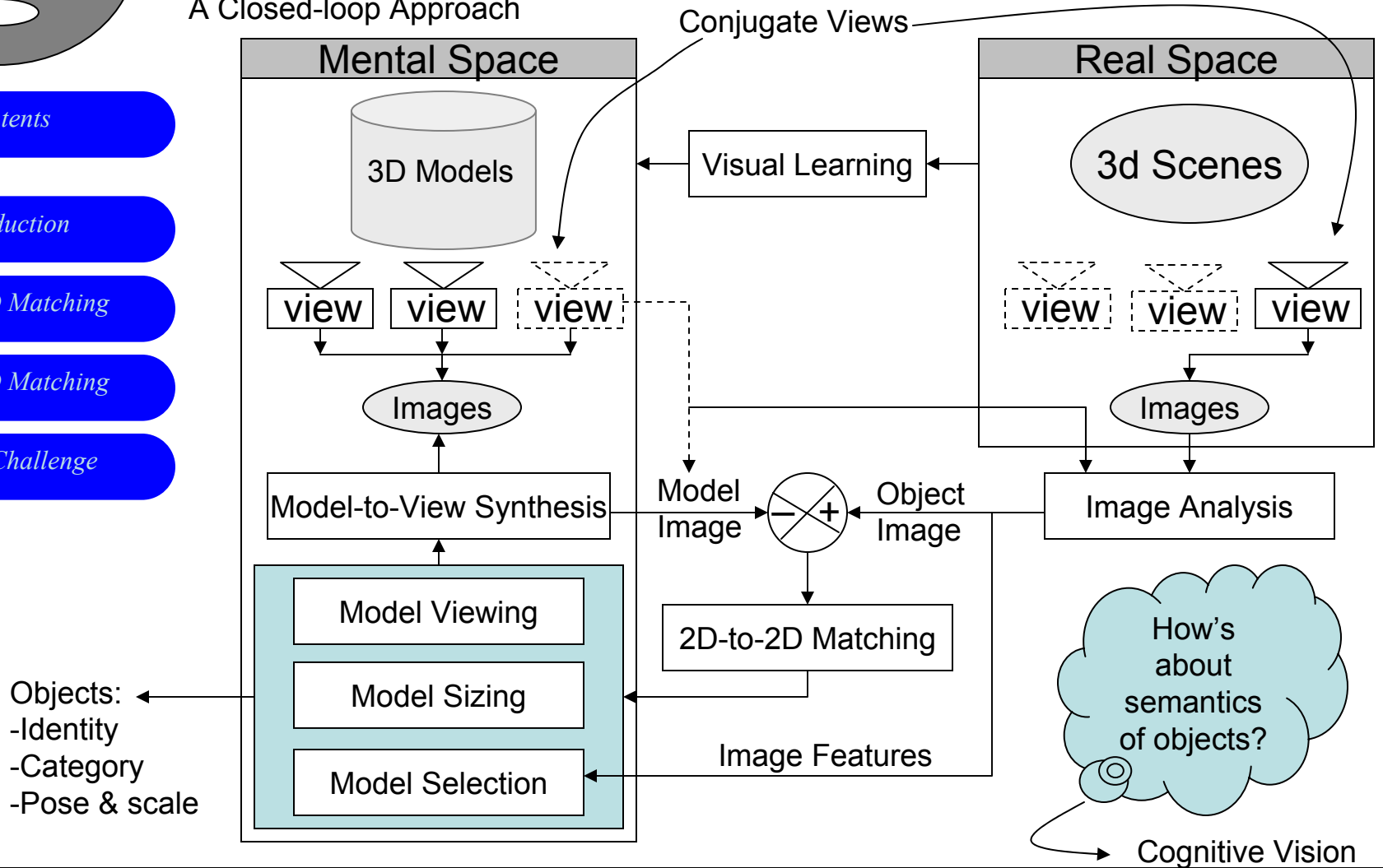
- We believe that it is worth investigating new engineering solutions which enable future robots to meaningfully interpret images, and faithfully recognize familiar objects.
- Here, we advocate a closed-loop approach for model-based vision, because:
 - Closed-loop is more robust than open-loop.
 - Most importantly, it is technically possible to bring the issue of inferring 3-D objects from 2-D images (or photos) down to the problem of 2D-to-2D matching.
 - In theory, it makes sense to consider 2D-to-2D matching, which is less difficult to tackle.



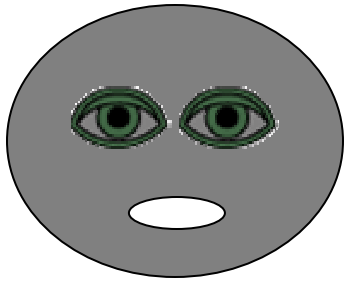
Future Challenge

- Contents
- Introduction
- 3D-to-3D Matching
- 2D-to-3D Matching
- Future Challenge

A Closed-loop Approach



Objects:
-Identity
-Category
-Pose & scale



Contents

Introduction

3D-to-3D Matching

2D-to-3D Matching

Future Challenge

End of Model-based Vision