AI in Computer Vision

Past, Present and Future

Image courtesy Amblin Entertainment

Ryan Wade
SBUID# 10557984
CSE 352 – Artificial Intelligence
Anita Wasilewska

Sources

- Shah, Mubarak. "Guest Introduction: The Changing Shape of Computer Vision in the Twenty-First Century." International Journal of Computer Vision 50 (2002): 103-110.
- Malisiewicz, Tomasz, and Alexei A. Efros. The Robotics Institute, Carnegie Mellon Universtity. Recognition by Association via Learning Per-exemplar Distances. June 2008. 3 Dec. 2008 <http://www.cs.cmu.edu/~tmalisie/projects/cvpr08/malisiewicz_cvp r08.pdf>.
- Fisher, Robert B. CVonline - Compendium of Computer Vision. School of Informatics, University of Edinburgh. 3 Dec. 2008 <http://homepages.inf.ed.ac.uk/rbf/cvonline/>.
- "Computer vision." Wikipedia. Computer Vision. 3 Dec. 2008 <http://en.wikipedia.org/wiki/computer_vision>.
- "ViPR Technology." Evolution Robotics, Inc. 3 Dec. 2008 <http://www.evolution.com/core/vipr.pdf>.

Overview

- What is Computer Vision?
- Early Computer Vision research
  - The MIT “Copy Demo”
- What happened During the Interim?
- Current Computer Vision research
- Current Computer Vision applications
- What does the future hold?

What is Computer Vision?

- The original goal of computer vision was to understand a single image of a scene, locate and identify objects, determine their structures, spatial arrangements, relationship with other objects, etc. (Shah, 2002)

The MIT “Copy Demo”

- 1970
- have a computer vision program analyze an image of a scene containing several stacked blocks, recover the structure of the blocks, and generate a script for a robot to build an exact copy of the block structure.

The MIT “Copy Demo”

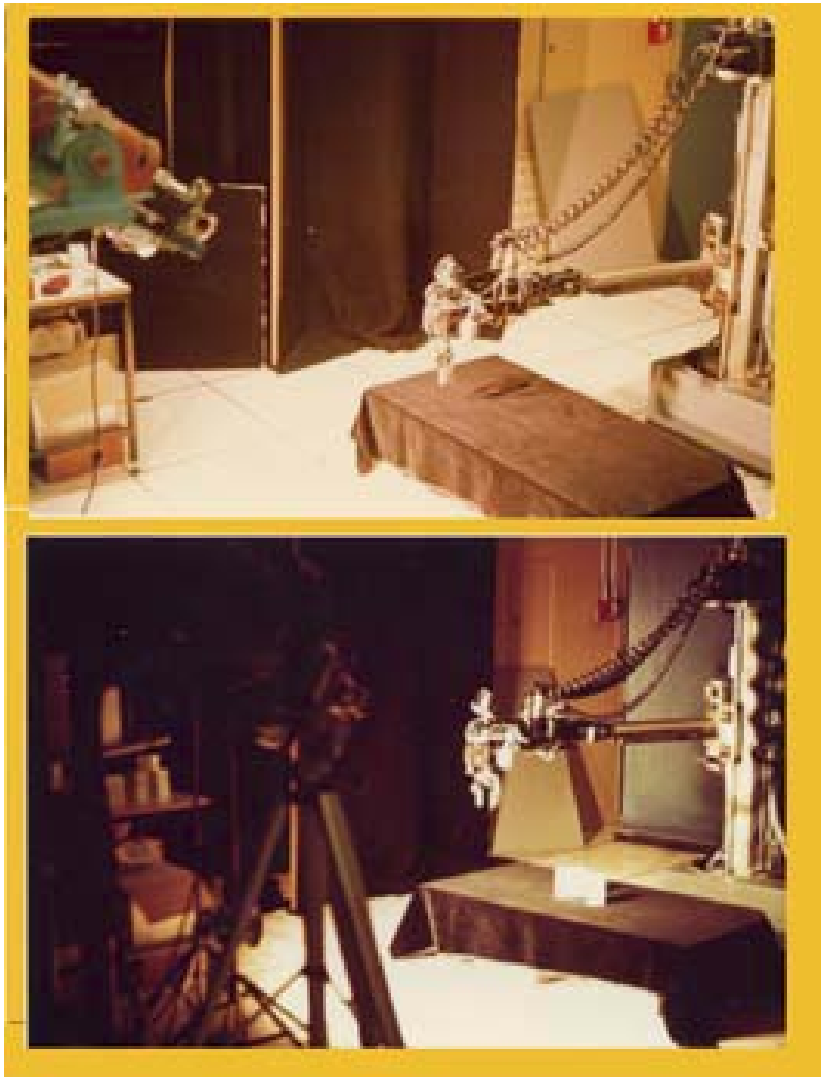


Image courtesy of Mubarak Shah

The MIT “Copy Demo”

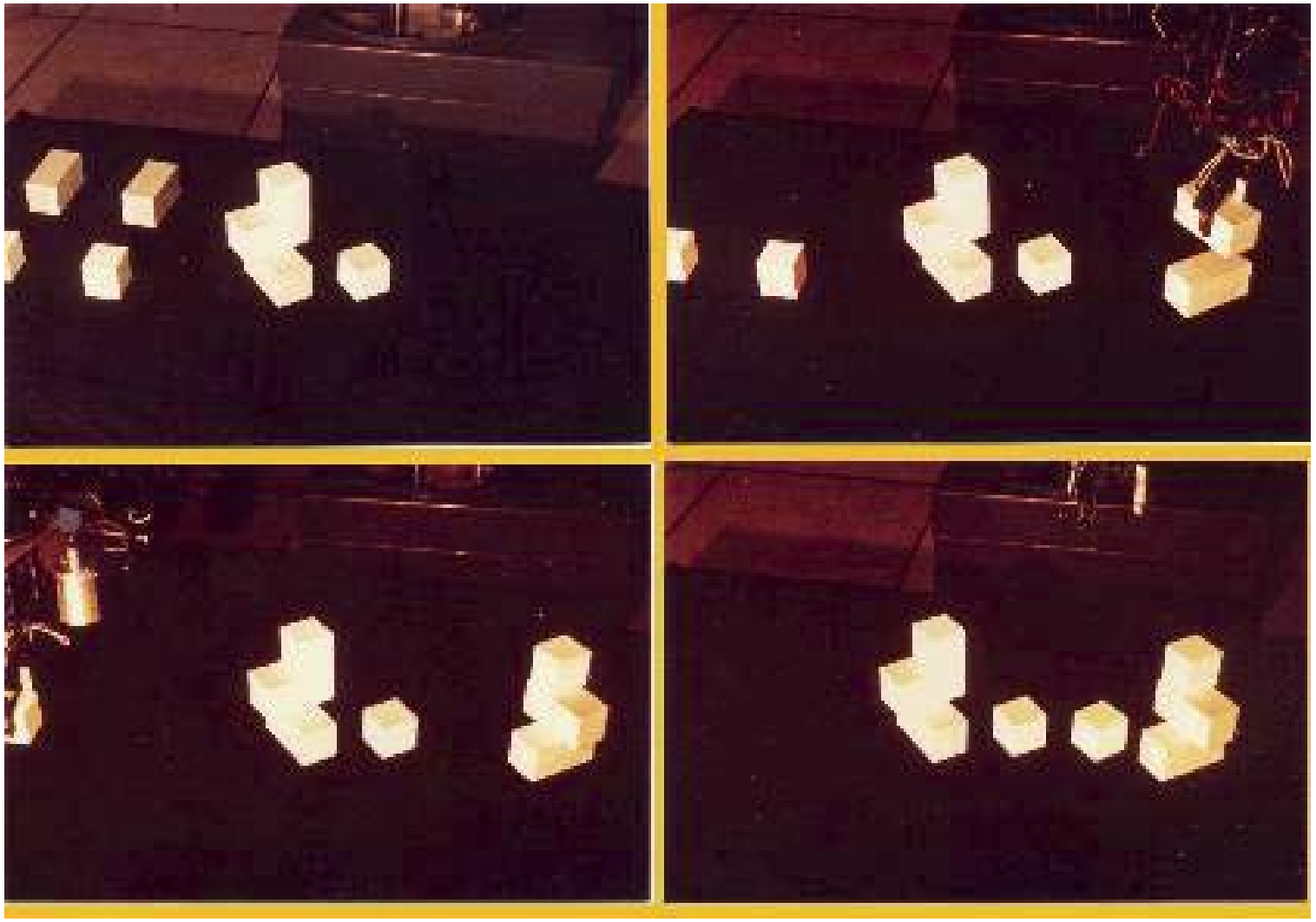


Image courtesy of Mubarak Shah

What did they find out?

- This was a high level vision problem; the researchers were not even able to extract lines from the image
- Low level vision (extraction of symbolic information) was not robust enough
- It became necessary to work from the ground-up, learning how to process images before trying to extract data from them

What happened next?

- Researchers worked on Low-level vision problems for a while (finding edges, trying to extract 3-D data from a 2-D image, etc.)
- Known as the “Mathematical Era” of computer vision
- Not relevant to CSE 352

Current Computer Vision Research

- Objects can be found in images, now research has moved onto interpretation of the objects (high-level computer vision)
- What is it, and what is it doing?
- Not only still images, but video as well
  - gesture recognition, activity description, facial expressions, etc.

Recognition by Association

- Don't ask "What is it," but rather "What is it *like*?"

Image courtesy of Tomasz Malisiewicz and Alexei A. Efros

Recognition by Association

- Recognition is usually assumed to mean “object naming” or categorization
- Categorizing objects is highly subjective
  - Some objects belong to same category, but look entirely different
  - Other objects look the same, but are assigned different categories

  - Categories are language independent

Recognition by Association

- Initial dataset consists of over 5000 images, segmented into 12,905 objects with 171 unique labels
- Each object is characterized by 14 different features (different attributes of shape, texture, color and location)
- New objects from unlabeled images are segmented and compared to the initial dataset
- Results as of June 2008 are promising, with roughly 80% accuracy
- Their goal is a system that uses unsupervised learning to recognize objects in images

Current Computer Vision Applications

- School of Informatics, University of Edinburgh lists 28 categories for Computer Vision, with a total of over 300 subcategories
- Includes:
  - Character Recognition (hand-written and printed)
  - Forensics
  - Human analysis (traits, behavior, identification)
  - Navigation, Mobility and Vehicle Control
  - Security
  - Robotics
  - Virtual Reality

So, when can we expect to see this kind of technology in our homes?

Simple visual recognition has already found its way into consumer products

Sony Aibo using ViPR Technology

Video courtesy of Evolution Robotics, Inc.

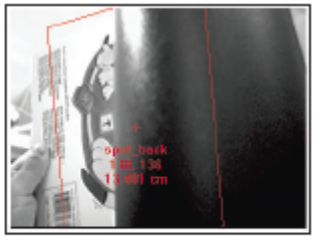
How does Aibo see?

- Unique features of a predetermined image are stored in a database
- An undisclosed algorithm is then used to find that image in the database

Recognition with partial occlusion

Recognition with different position

Recognition at a distance



Recognition at an angle with partial occlusion

Images courtesy of Evolution Robotics, Inc.

What does the future hold?

Video courtesy of Matt Groening

Any Questions?