What is Computer Vision?

Prof. Dr. Elli Angelopoulou

Pattern Recognition Lab (Computer Science 5)

University of Erlangen-Nuremberg

Introduction

- The goal of this presentation is to give a brief introduction and overview of the field of
  - Computer Vision
- An atypical computer science discipline
- Multidisciplinary
  - Programming
  - Algorithms
  - Geometry
  - Optics

Outline

- Definition
- Brief History
- Applications
- The importance of shape (geometry) and optics
- Brief overview of widely used computer vision techniques. Most of these topics we will cover in during the course of the semester.

What is Computer Vision?

- **Computer vision** involves the automatic deduction of the *structure* and the *properties* of a possibly dynamic three-dimensional world from either a single or multiple two-dimensional images of the world.

```
Example

Input: Image on the left

Output:

1 windmill: 3 stories tall, 4 blades
(1 hidden), conical roof;

5 people: 3 male, 2 female;

1 mill stone;

1 stone wall
```

How it all started

- The term Computer (Machine, Robot) Vision was first introduced as a special topic in Artificial Intelligence.
- First attempts: Tracing boundaries of polygonal objects.
- Revolutionary work by David Marr around 1975 at the Massachusetts Institute of Technology.
- First use of a pair of cameras for mimicking biological eyes in the 1960s

Computer Vision

- Computer Vision evolved as a stand-alone field around in the late 1970s
- Vision moved beyond “biological imitation” when it started being applied in factory automation as a robotic sensor (term Robot Vision started appearing)
- Different schools of thought:
  - Physics and math oriented
  - Statistical analysis
  - Neural networks
  - Heuristic approaches

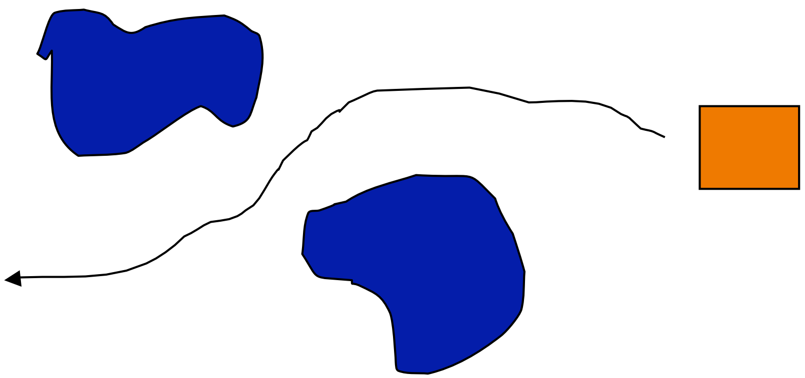
} LME

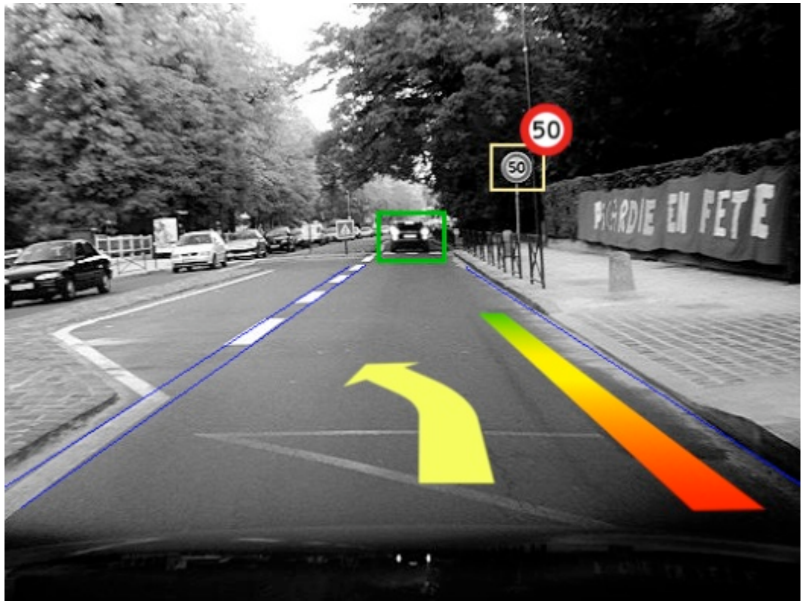
Applications

- Navigation (autonomous vehicles)
- Factory automation (assembly and packaging)
- Tele-presence (Telemedicine, virtual presence in museums, athletic events, like a basketball game)
- Object recognition (Automatic Target Recognition)
- Object tracking (surveillance)
- Human detection and identification (security and surveillance)
- Motion analysis (weather forecasting)
- Image retrieval (database or web-page search)

The Role of Computer Vision

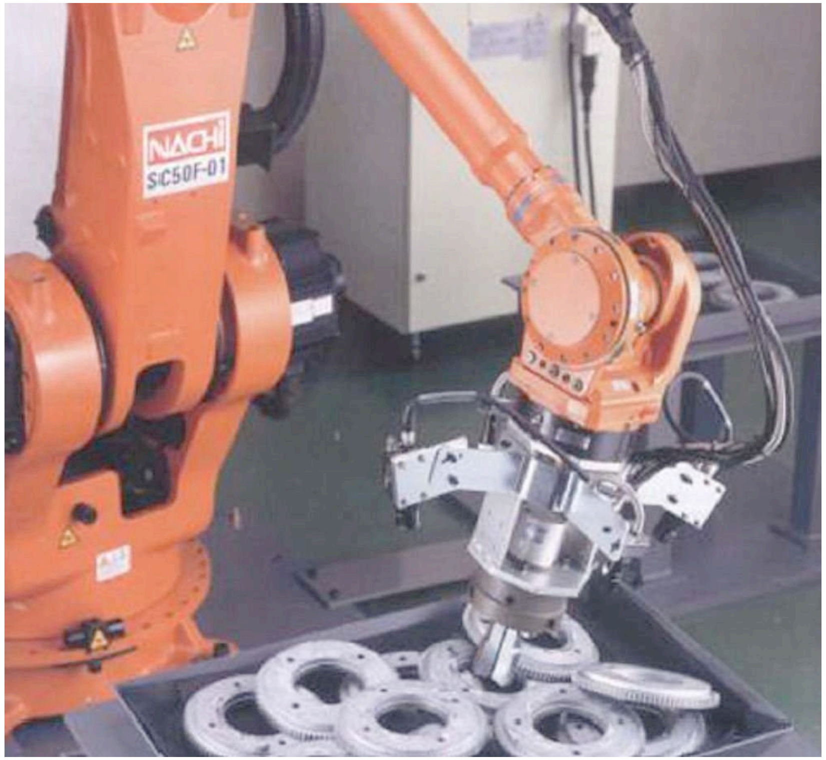
- Navigation
  - Compute distance to the various obstacles
  - Compute path that guarantees shortest safe path
  - Identify different types of objects in its path (people, cars, roadsigns, etc.)





The Role of Computer Vision

- Factory Automation
  - Identify object to be manipulated
  - Compute its shape, color or other properties
  - Quality assessment
  - Compute shortest and safest trajectory of robotic grasping arm



The Role of Computer Vision

- Tele-presence
  - Compute the dimensions, shape and location of each object in the different locations.
  - Merge the scenes in one virtual scene that is geometrically correct (proper locations, not overlapping)
  - Merge the scenes in one virtual scene that is optically correct (shadows, inter-reflections, same background, consistent lighting)

The Role of Computer Vision

- Object Recognition (initial work focused a lot on Automatic Target Recognition -ATR)
  - Compute dimensions of objects
  - Classify objects as possible targets
  - Compute location of each possible target and/or trajectory to it.

Original Scene

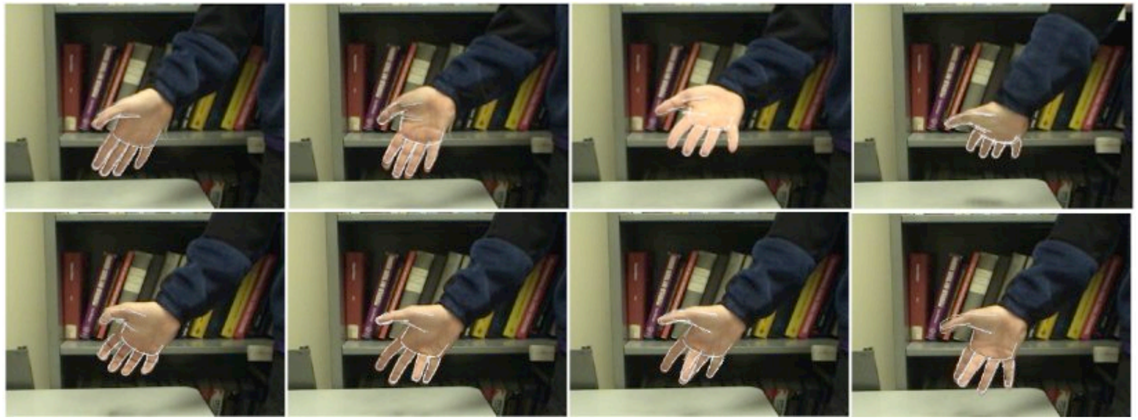
Nominated Targets Based

From the CFP of the *ECCV Workshop on Parts and Attributes* 2010

The Role of Computer Vision

In a sequence of images taken over a period of time

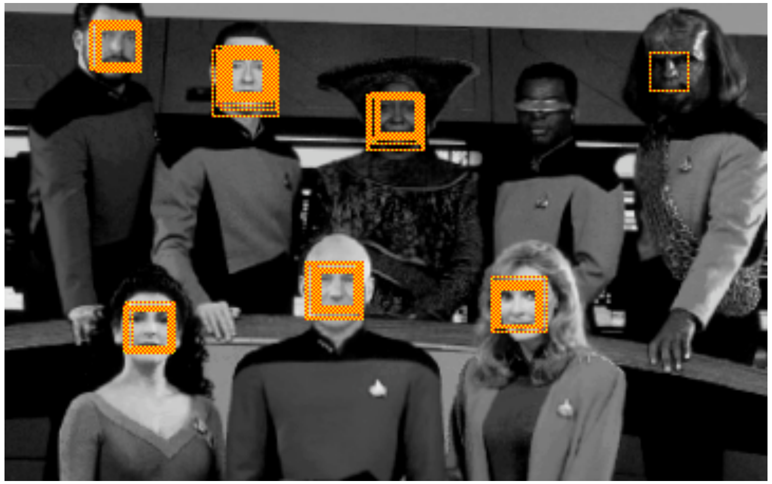
- Object Tracking
  - Identify the object of interest
  - Compute its location at each time instance *t*.
- Motion Analysis
  - Identify which objects are moving in the scene
  - Compute their velocity



"Visual Hand Tracking Using Occlusion Compensated Message Passing" by Erik B. Sudderth, Michael I. Mandel, William T. Freeman and Alan S. Willsky.

The Role of Computer Vision

- Human Detection and Identification
  - Compute the location of faces in a cluttered scene
  - Identify a specific individual under varying conditions

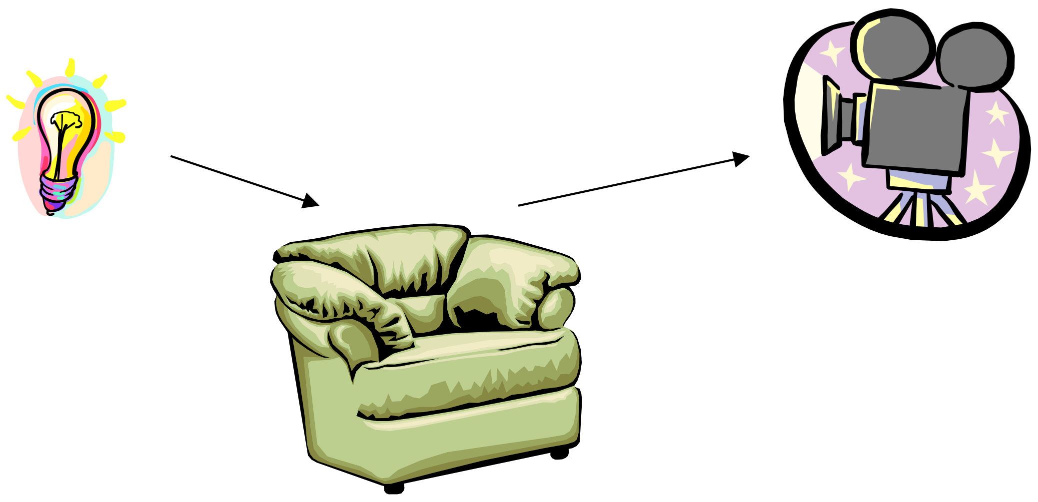


Bottom Line

- The majority of applications involve the (ideally robust) computation of a quantitative description of the objects in the captured scene.
- Quantitative description
  - geometry (shape) of objects in the scene
  - material, color or other properties of the objects in the scene
  - persistence in measurements independent of viewing conditions
- Reverse engineer the process that caused the image to be formed.
- Semantic gap
  - go beyond quantitative analysis
  - extract more abstract descriptions (chair, table, painting, upset person, lost/forgotten item)

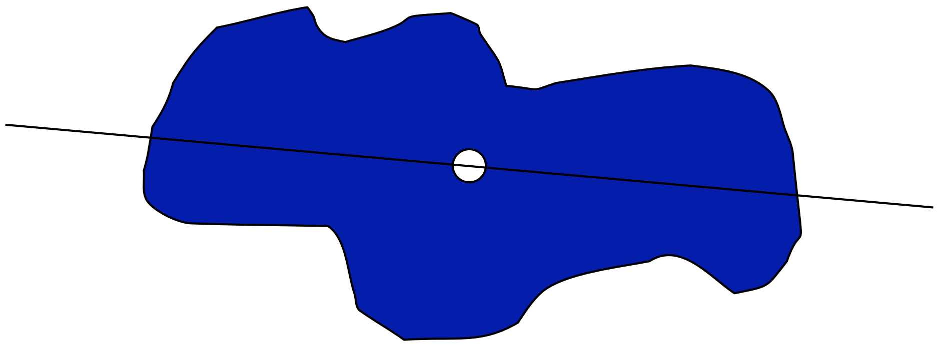
Image Formation

- There are three major components that determine the appearance of an image
  - Geometry
  - Optical properties of the materials in the scene
  - Illumination conditions



Basic Shape Analysis

- The center of black and white silhouettes can be easily computed using moment analysis
  - 0^{th} order moment → size
  - 1^{st} order moment → center of mass
  - 2^{nd} order moments → orientation information

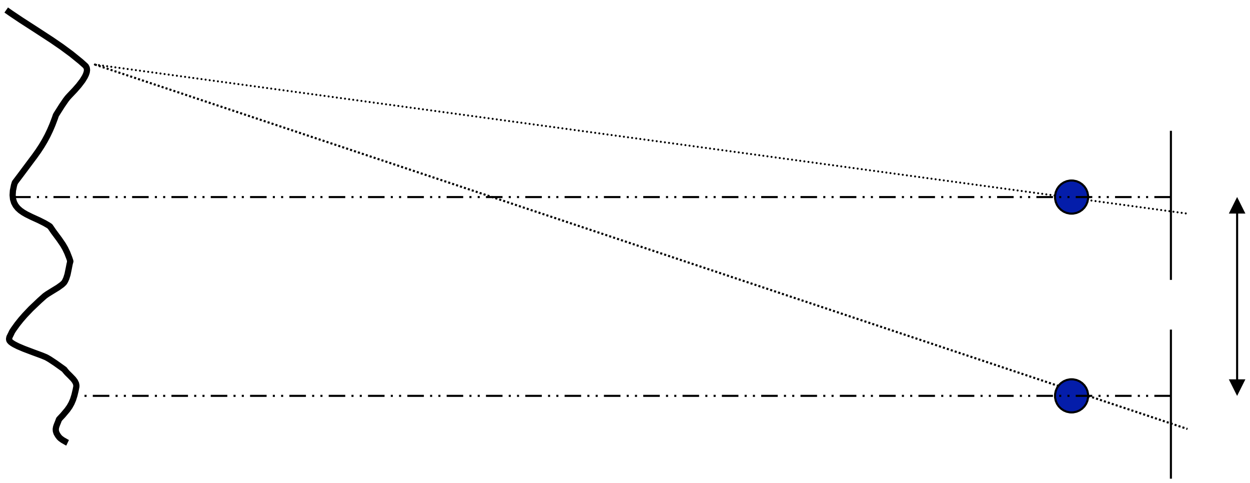


Extraction of Silhouettes

- Edge detection
- Biological evidence that animals perform some form of differentiation on the images
- Further analysis is done on 2.5 D sketch: 2D image formed on retina + edge information (Marr)

Depth Computation

- Binocular (poly-ocular stereo)
- The "shifting" of the scene between the 2 images provides the depth information



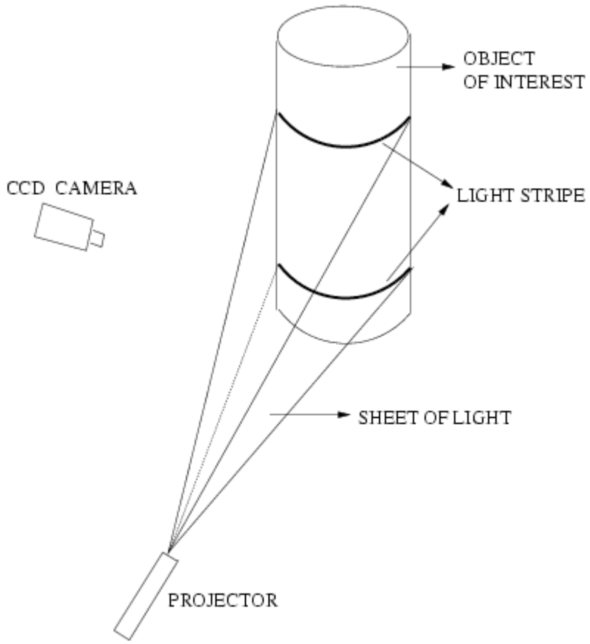
- What if there are not enough uniquely identifiable points?

Shape from Shading

- Shading provides shape clues (disk versus sphere)
- In the 1970s it was proved by Horn that the shape of a surface can be extracted from a single image, if we know how the surface is illuminated.
- Main idea:
  - The variations in shading of a single-colored object are caused by changes in the geometry of the object.
  - You are given the relationship between the shape of the object and the shading variations
  - A camera captures these shading variations
  - Extract the geometry

Structured Light

- Project a light beam of known geometry (e.g. a collection of thin vertical stripes) onto a scene
- Take a picture of the scene illuminated by the structured light
- The shape of the objects on the scene distorts the light pattern. Use that distortion to deduce the shape of the object



Motion Analysis

- Main idea: Track features as they move from one frame to the next
- A basic technique:
  - Extract edges at each frame of the movie
  - Compute the motion of these edges in the 2D frames
  - Relate 2D motion in image with 3D motion
- What happens if the scene changes abruptly? (lights are turned off)
- Does the shadow of moving clouds get interpreted as motion, when there shouldn't be any?

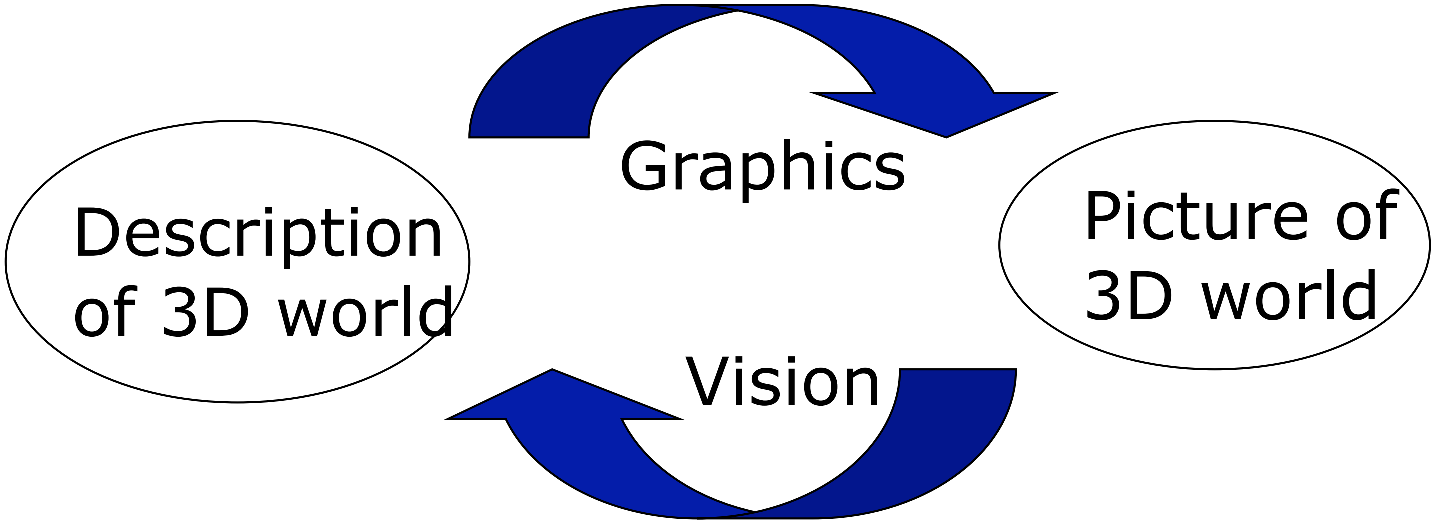
Shape Analysis

- Extract invariant shape descriptors that can be used in object recognition
- Ideally descriptors should be succinct to facilitate information transmission
- Example: Curvature information

Computer Vision vs. Image Processing

- Image processing typically deals with the early processing stages.
- Conversion of sensed light into an image file
- Noise removal
- Image enhancement
- Image compression
- Typically, the input is an image and the output is also an image
- Treats the input as a signal

Computer Vision vs. Computer Graphics

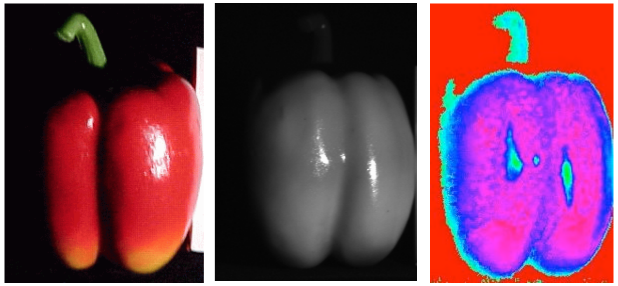


Shared Tools: underlying theory (optics, geometry)
algorithms

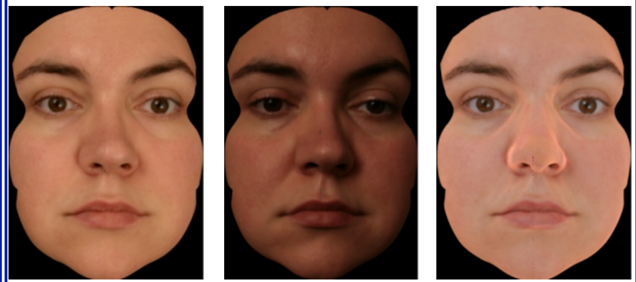
Computer Vision vs. Medical Imaging

- Medical Imaging was originally part of Computer Vision
- Different imaging modalities with very distinct image formation processes.
- More constrained set of objects that appear in medical images (easier to use prior knowledge).
- High demands in accuracy.

Computer Vision - Research Projects



Reflectance analysis



Skin reflectance

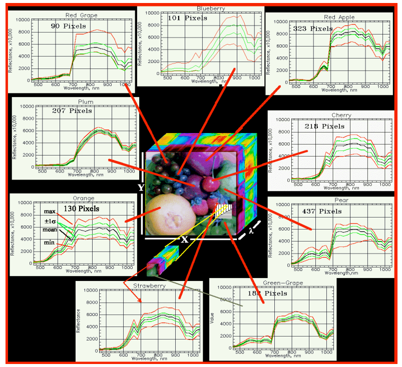
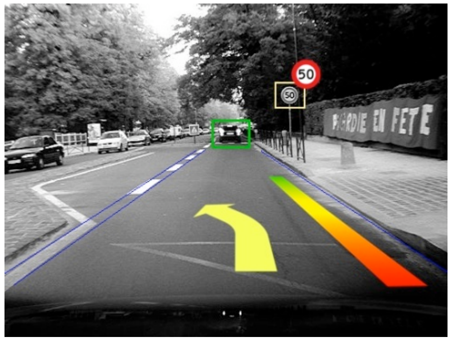


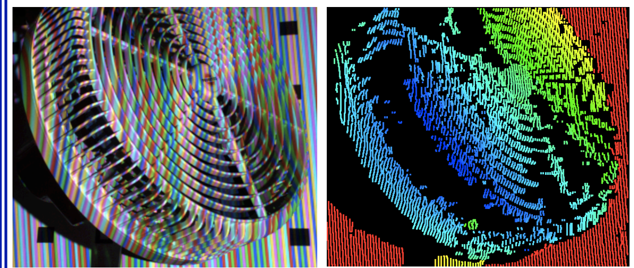
Image courtesy of Dr. Nahum Gat, OKSI, Inc.

Multispectral imaging

Image forensics



Context-aware navigation



3D Reconstruction

Summary

- Computer Vision is a multidisciplinary field.
- Many diverse topics.
- In order to be able to apply oneself in computer vision one must have an understanding of:
    - Image formation process
    - Basic image processing methods
    - Information that can be extracted from single images
    - Combination of information from multiple images
    - Implementation of algorithms (real time issues, accuracy issues etc.)
- Upon completion of the class, one should:
    - Have a good understanding of the aforementioned topics
    - Be able to formally argue about the effectiveness a computer vision system, and implement and test a prototype.