Computer Vision / Digital Image Processing
Introduction

Dr. Muhammad Usman Ghani
Associate Professor
CS & Engg. Dept. UET lahore

January 31, 2017

Outline

Agenda

Today Discussion

- Definition
- Image Processing overview
- Image Processing applications
- Key stages in Image Processing
- Some term projects examples

Readings (latest editions from market)

- Gonzalez, R.C. Digital image processing.
- Gonzalez, R.C. Digital image processing using Matlab.

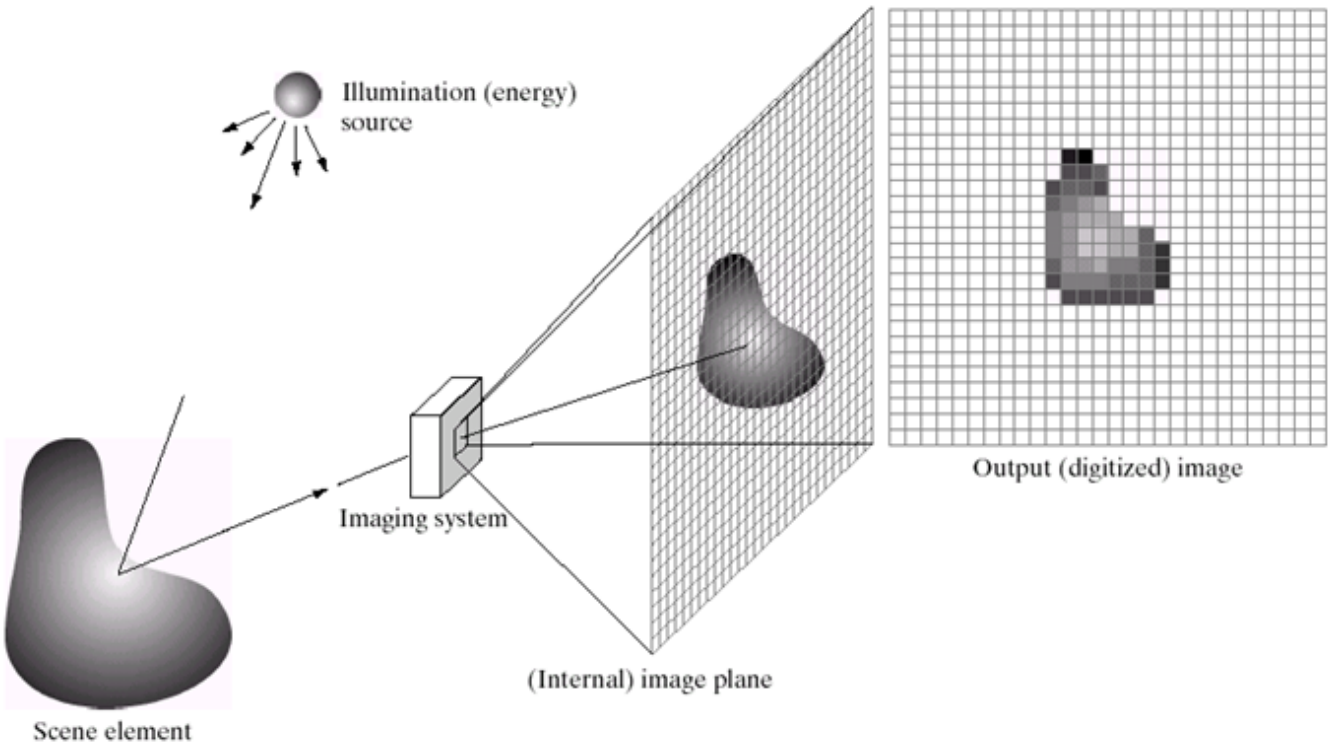
Introduction

One picture is worth more than ten thousand words (Anonymous)

Definition

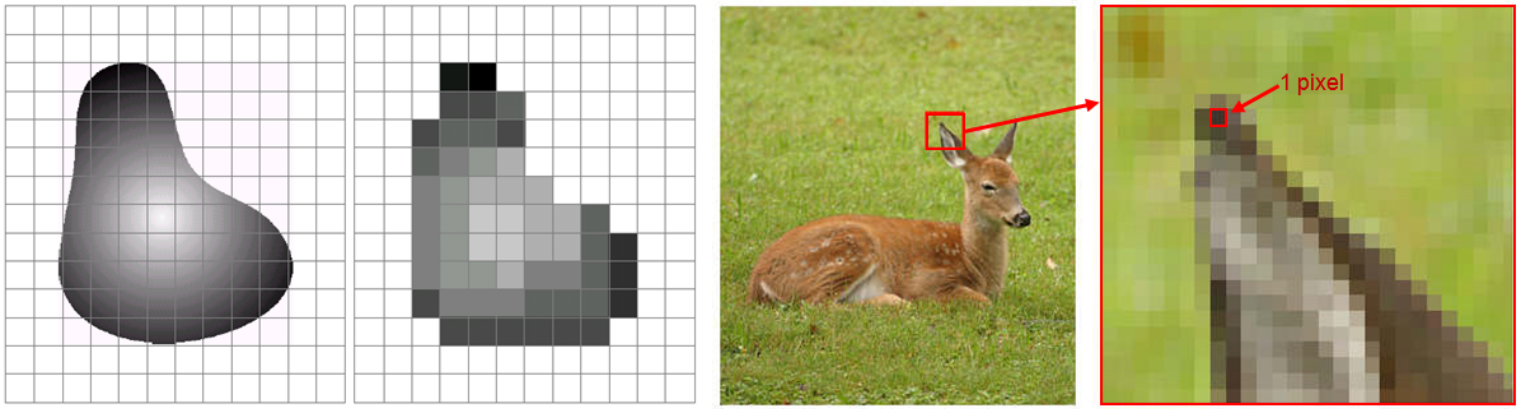
* Vision is about discovering from images **what** is present in the scene and **where** it is.
* In Computer Vision a camera (or several cameras) is linked to a computer. The computer interprets images of a real scene to obtain information useful for tasks such as navigation, manipulation and recognition.

Definitions Digital Image



- A digital image is a representation of a two-dimensional image as a finite set of digital values, called picture elements or pixels

Definitions Pixels



- Pixel values typically represent gray levels, colours, heights, opacities etc
- Remember digitization implies that a digital image is an approximation of a real scene

Goal of Digital Image Processing

- Digital image processing focuses on two major tasks
  - Improvement of pictorial information for human interpretation
  - Processing of image data for storage, transmission and representation for autonomous machine perception.

Some argument about where image processing ends and fields such as image analysis and computer vision start

Goal of Computer Vision

To bridge the gap between pixels and meaning

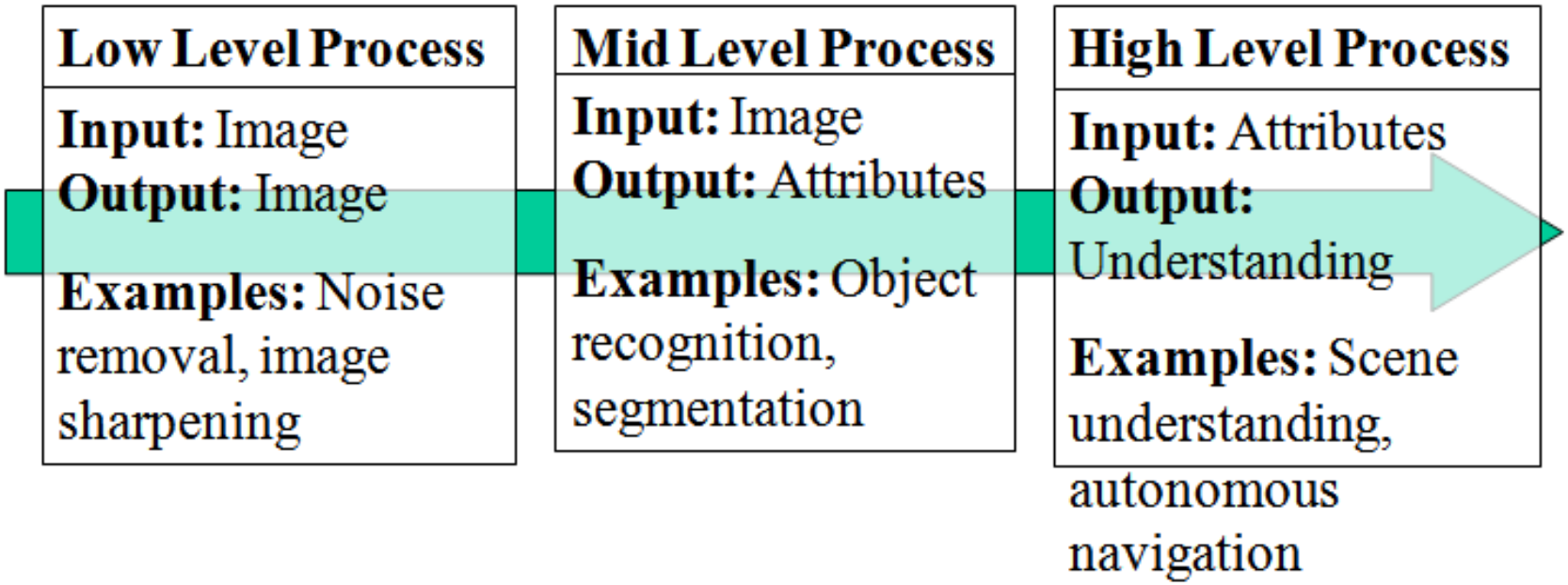
0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

Figure: What we see

What computers see

DIP to CV

The continuum from image processing to computer vision can be broken up into low-, mid- and high-level processes



Human vs Machine Vision

Can computers match (or beat) human vision?

- Yes and no (but mostly no!)
  - humans are much better at "hard" things
  - computers can be better at "easy" things

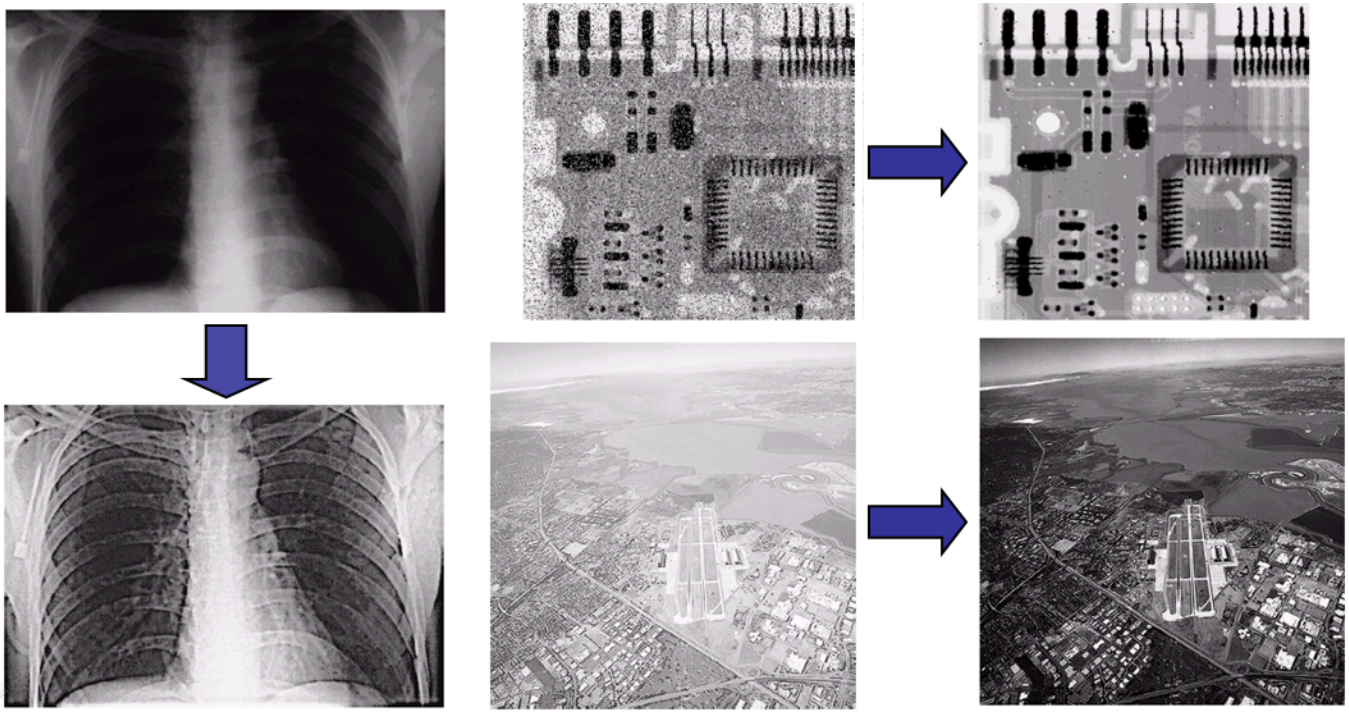
Human vs Machine Vision

Human perception has its shortcomings

Figure: Example where humans make mistakes that computers can avoid

Examples: Image Enhancement

One of the most common uses of DIP techniques: improve quality, remove noise etc



Examples: Artistic Effects

Artistic effects are used to make images

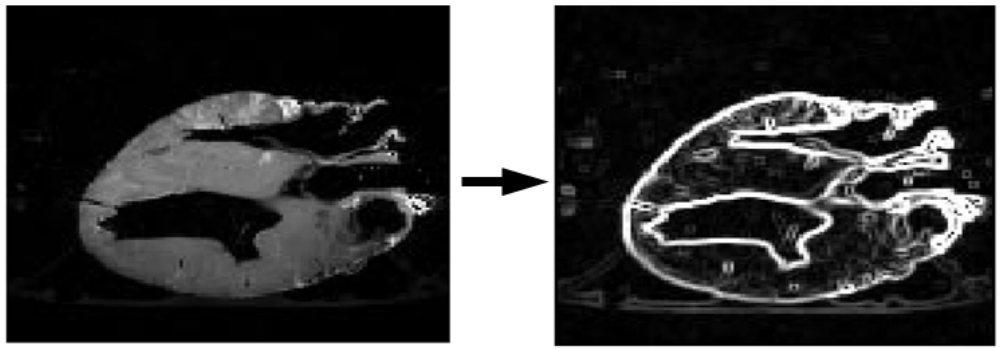
- more visually appealing
- to add special effects and
- to make composite images




Examples: Medicine

Take slice from MRI scan of canine heart, and find boundaries between types of tissue

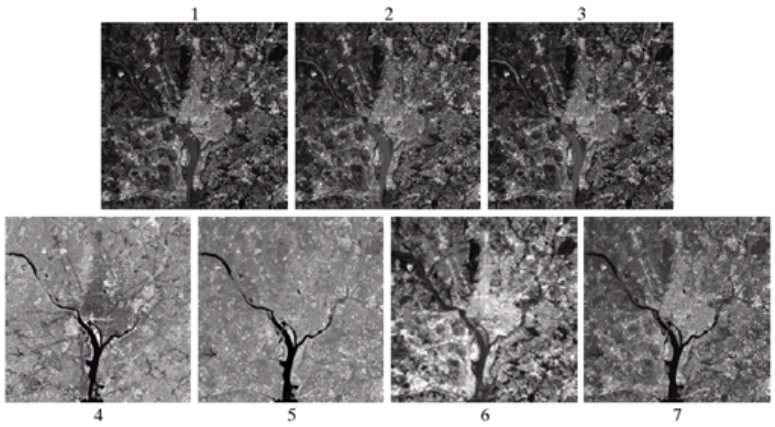
- Image with gray levels representing tissue density
- Use a suitable filter to highlight edges

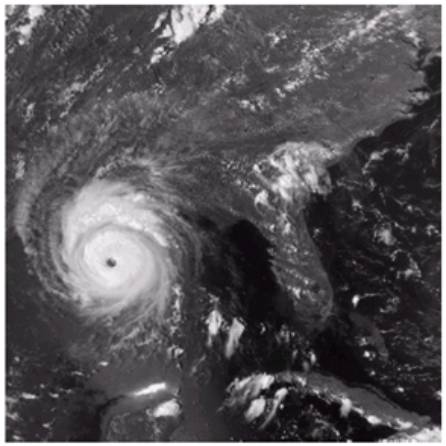


Examples: GIS

Geo-graphic information systems

- Digital image processing techniques are used extensively to manipulate satellite imagery
- Terrain classification
- Meteorology

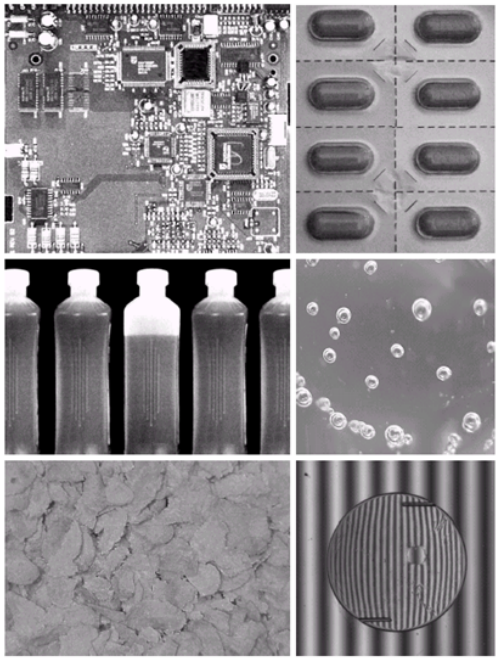




Examples: Industrial Inspection

Human operators are expensive, slow and unreliable

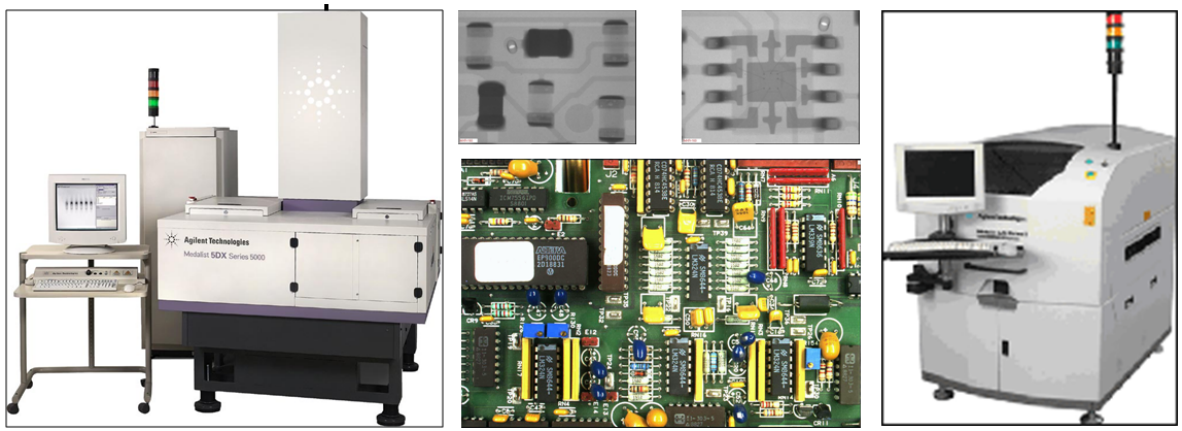
- Make machines do the job instead
- Industrial vision systems are used in all kinds of industries
- Can we trust them?



Examples: PCB Inspection

Printed Circuit Board (PCB) inspection

- Machine inspection is used to determine that all components are present and that all solder joints are acceptable
- Both conventional imaging and x-ray imaging



State of the art

The next slides show some examples of what current vision systems can do

Earth viewers (3D modeling)

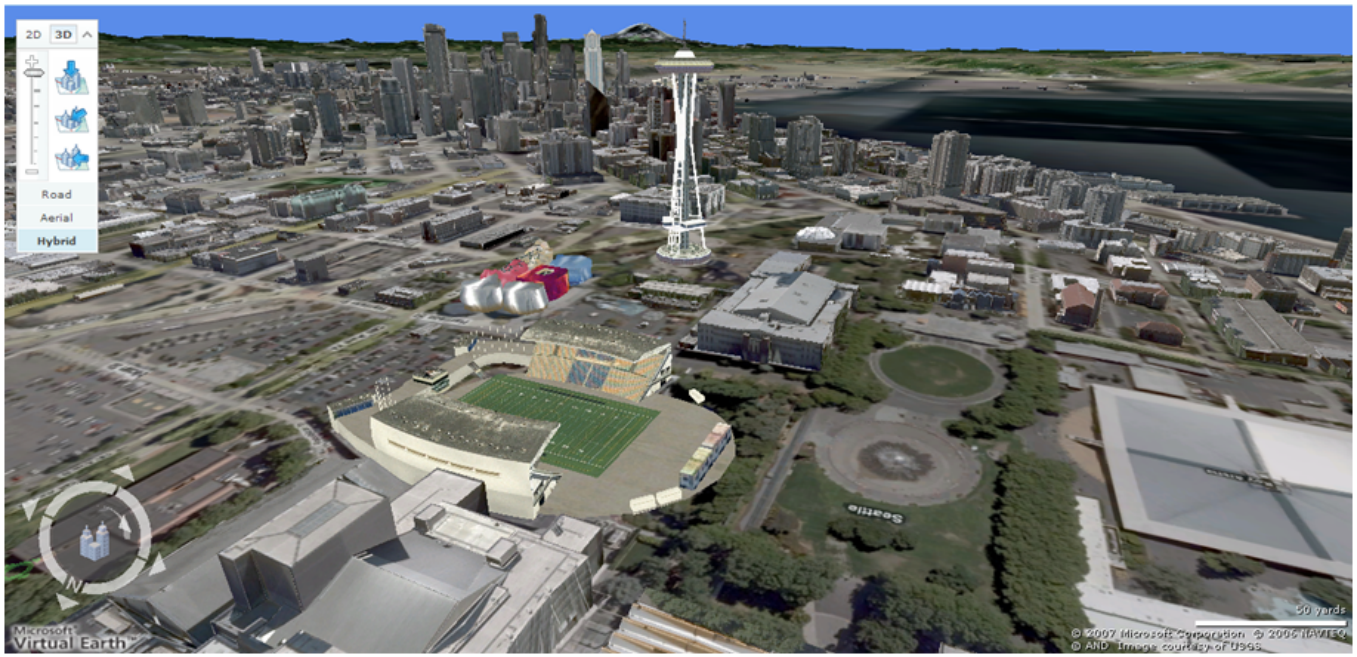


Figure: images from virtual earth [1]

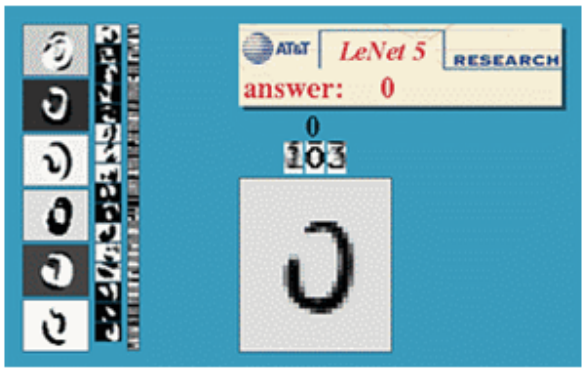
Google maps are common these days.

[1] http://www.microsoft.com/maps/

Optical character recognition (OCR)

Technology to convert scanned docs to text

- If you have a scanner, it probably came with OCR software



Digit recognition, AT&T labs
http://www.research.att.com/~yann/

License plate readers
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

Face detection

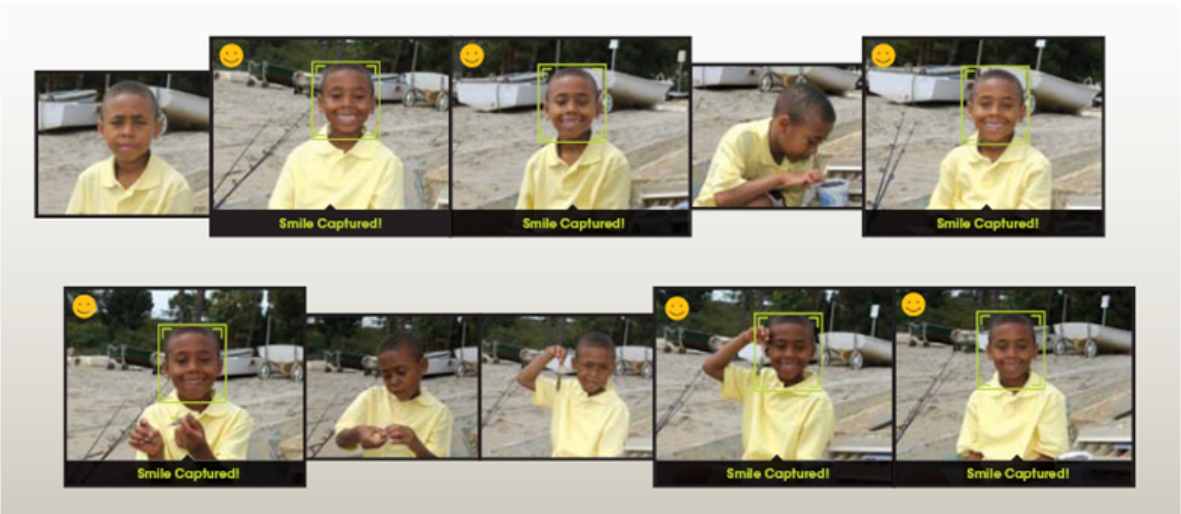
Many new digital cameras now detect faces

- Canon, Sony, Fuji, *etc.*

Why is this useful?

- Main reason is focus. Also enables smart cropping.

Smile Detection

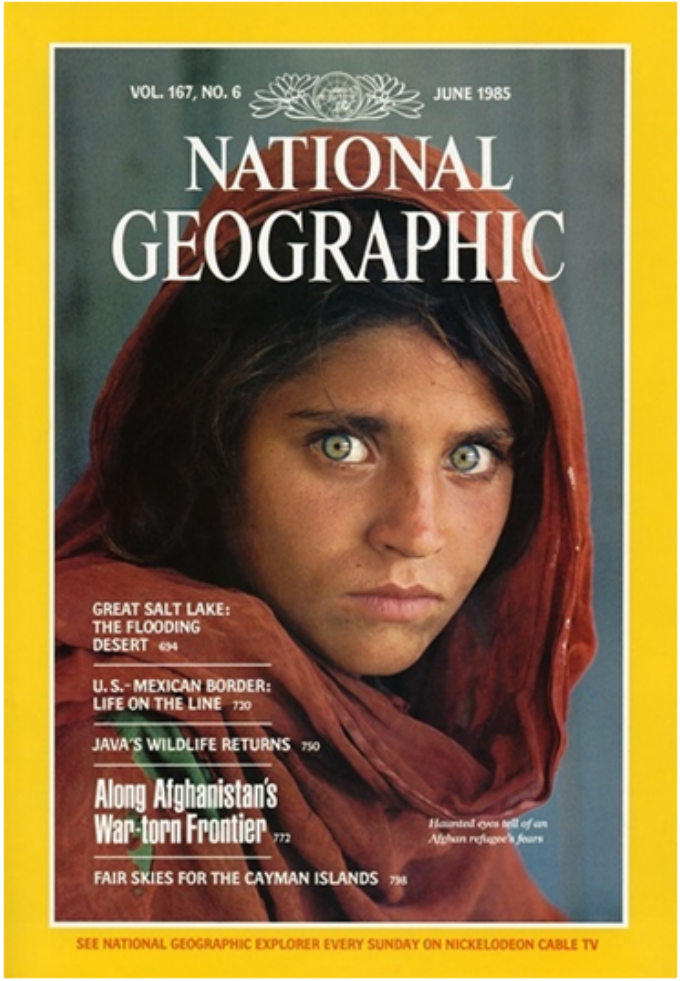


Object Recognition

LaneHawk by EvolutionRobotics

- A smart camera is flush-mounted in the checkout lane, continuously watching for items. When an item is detected and recognized, the cashier verifies the quantity of items that were found under the basket, and continues to close the transaction. The item can remain under the basket, and with LaneHawk,you are assured to get paid for it

Face Recognition



Vision-based biometrics

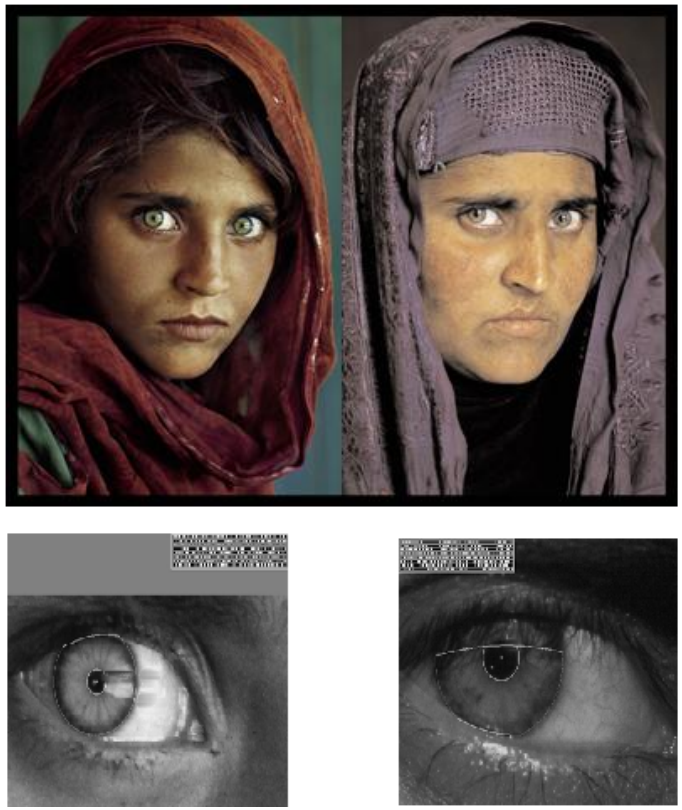


Figure: How the Afghan Girl was Identified? — by Her Iris Patterns?

Login without a password

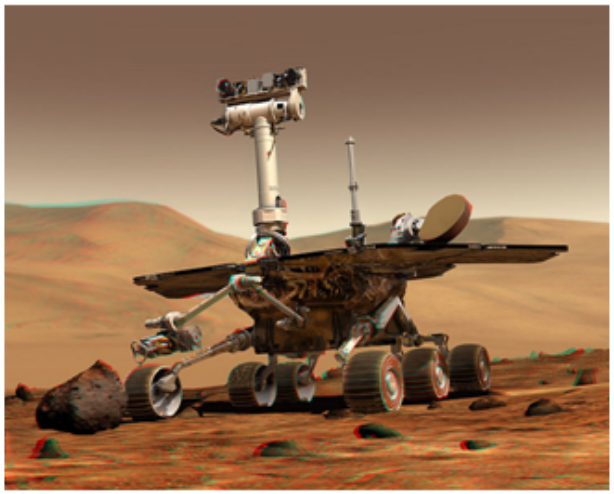
Figure: Fingerprints of thumbs, face recognition

Object recognition (in mobile phones)

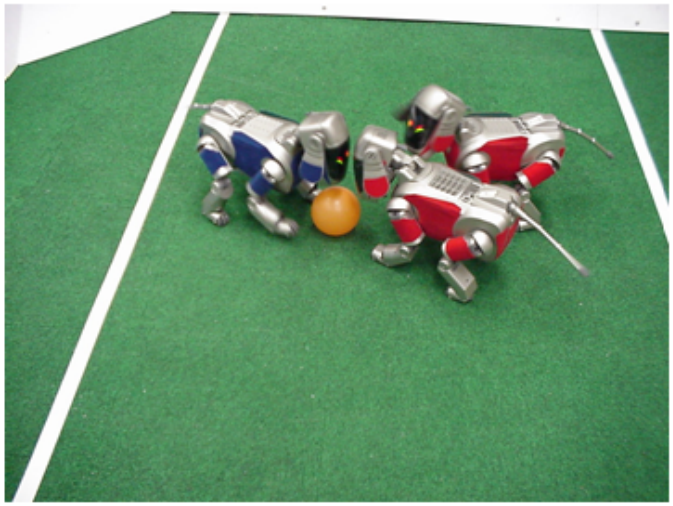
This is becoming real:

- Lincoln Microsoft Research
- Point & Find, Nokia

Robotics

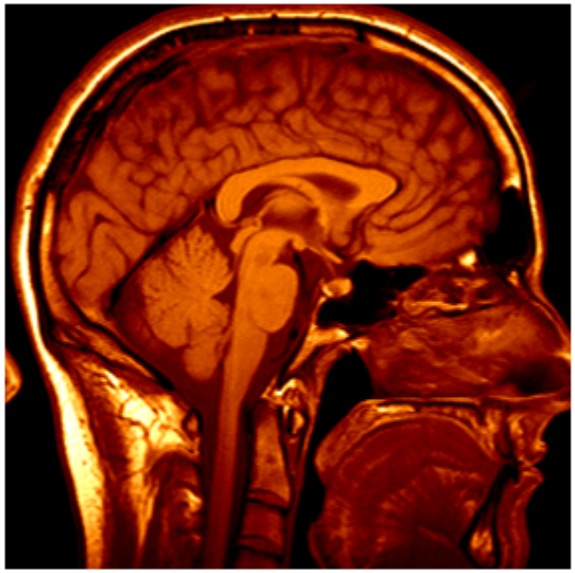


NASA's Mars Spirit Rover
http://en.wikipedia.org/wiki/Spirit_rover



http://www.robocup.org/

Medical Imaging



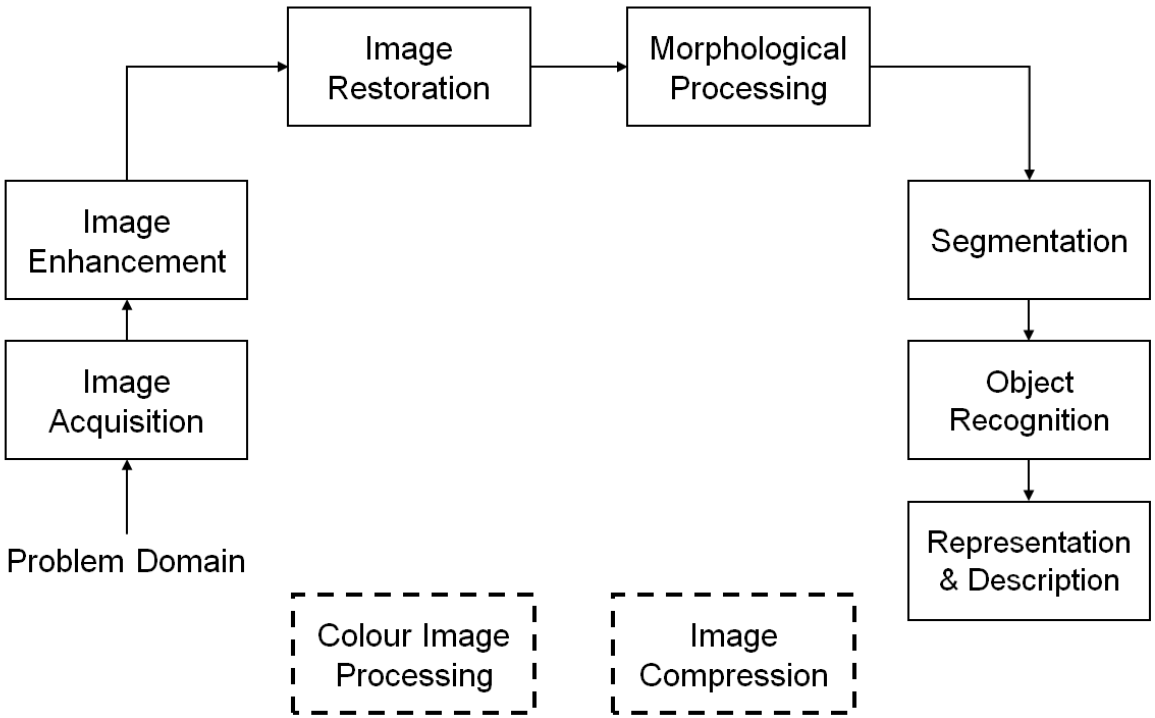
3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT

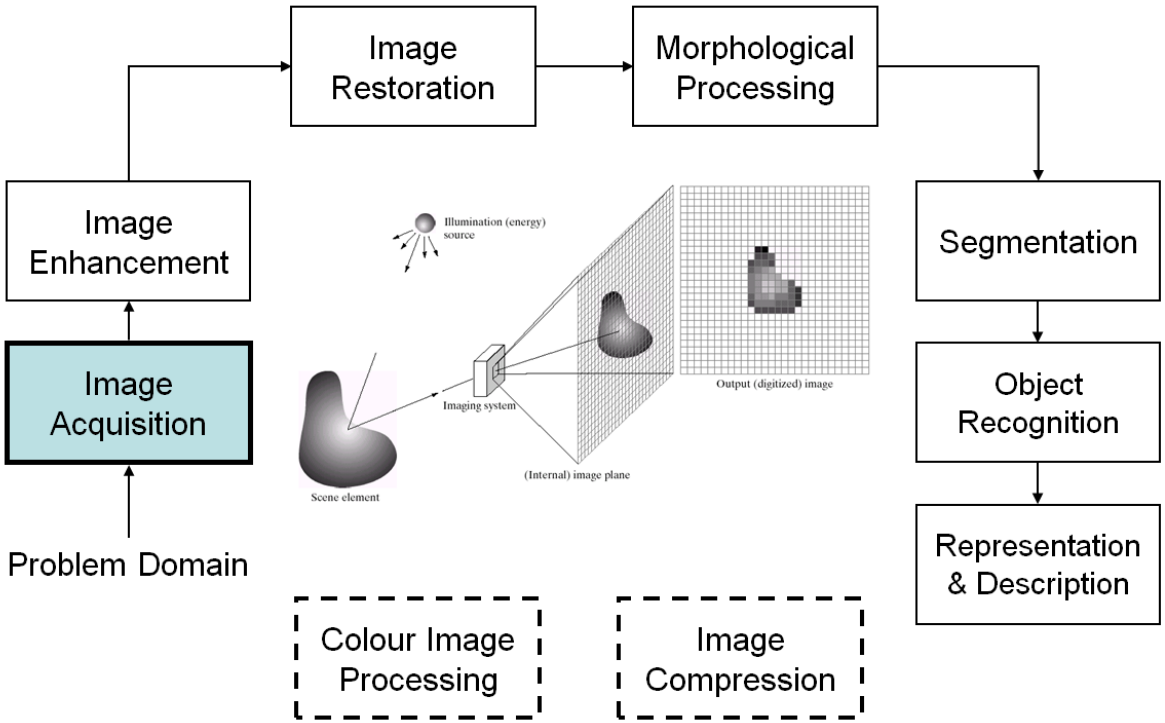
Current state of the art

- You just saw examples of current systems
  - Many of these are less than 5 years old
- This is a very active research area, and rapidly changing
  - Many new apps in the next 5 years
- To learn more about vision applications and companies
  - David Lowe maintains an excellent overview of vision companies
    + http://www.cs.ubc.ca/spider/lowe/vision.html

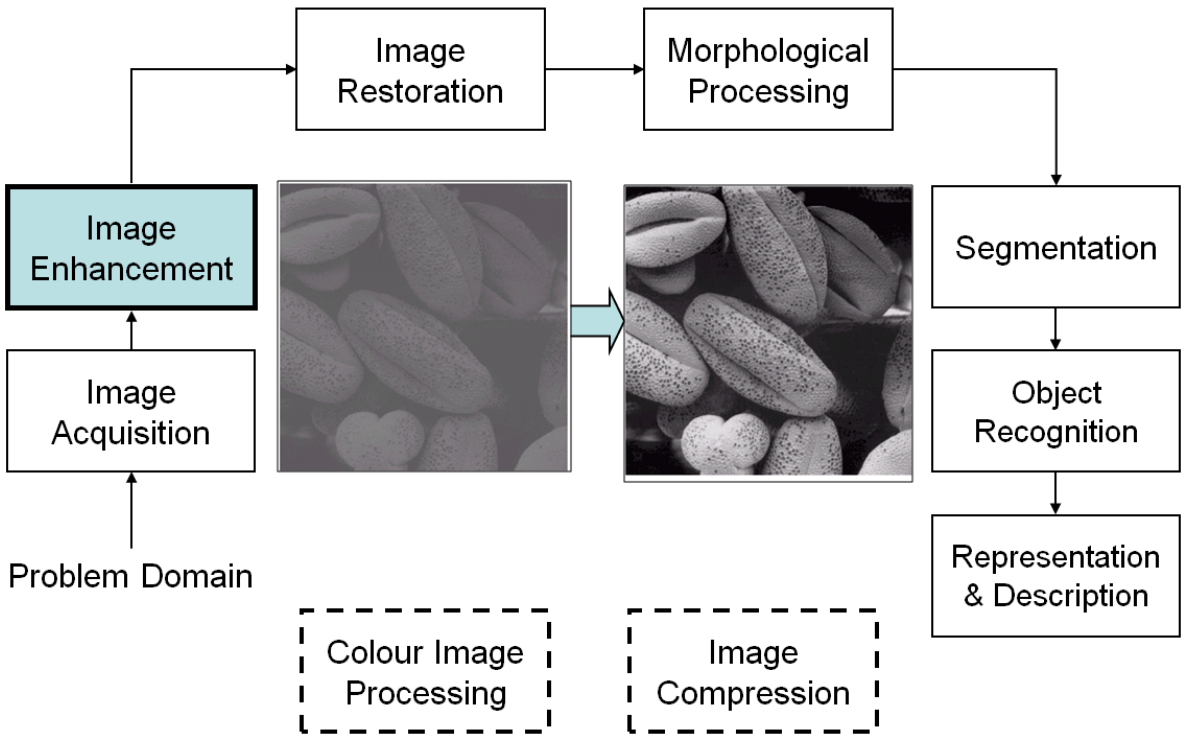
Key Stages in Digital Image Processing



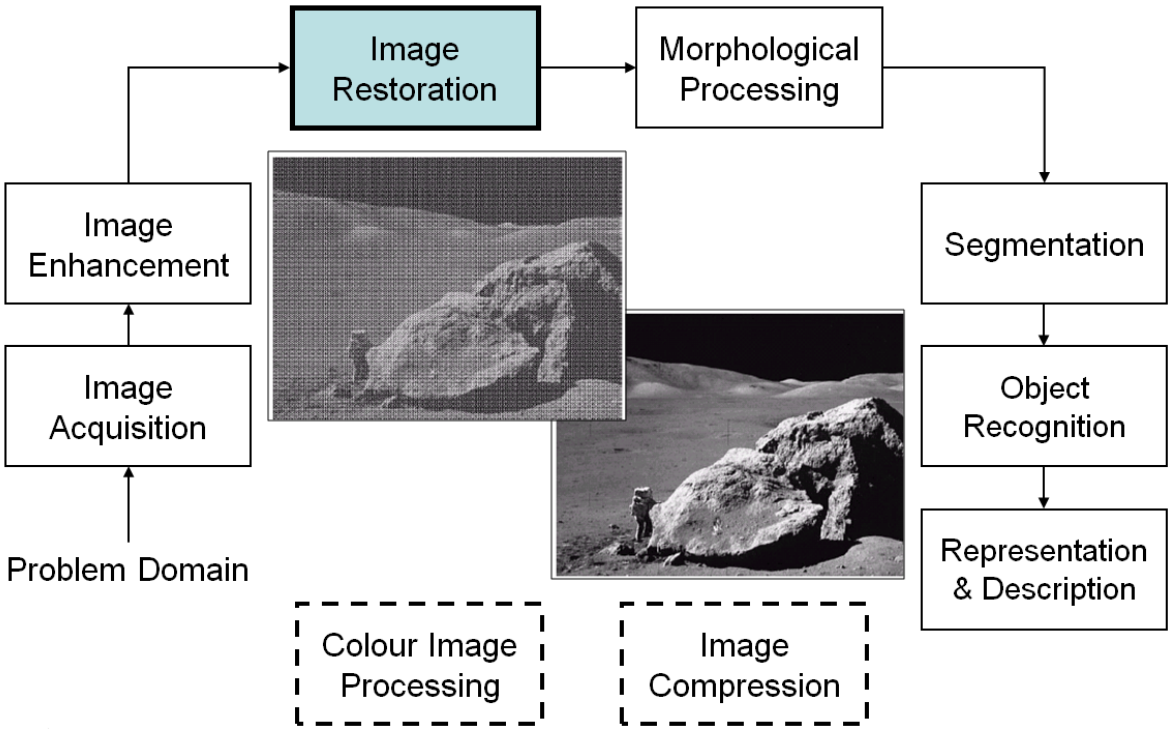
Key Stages in DIP: Image Acquisition



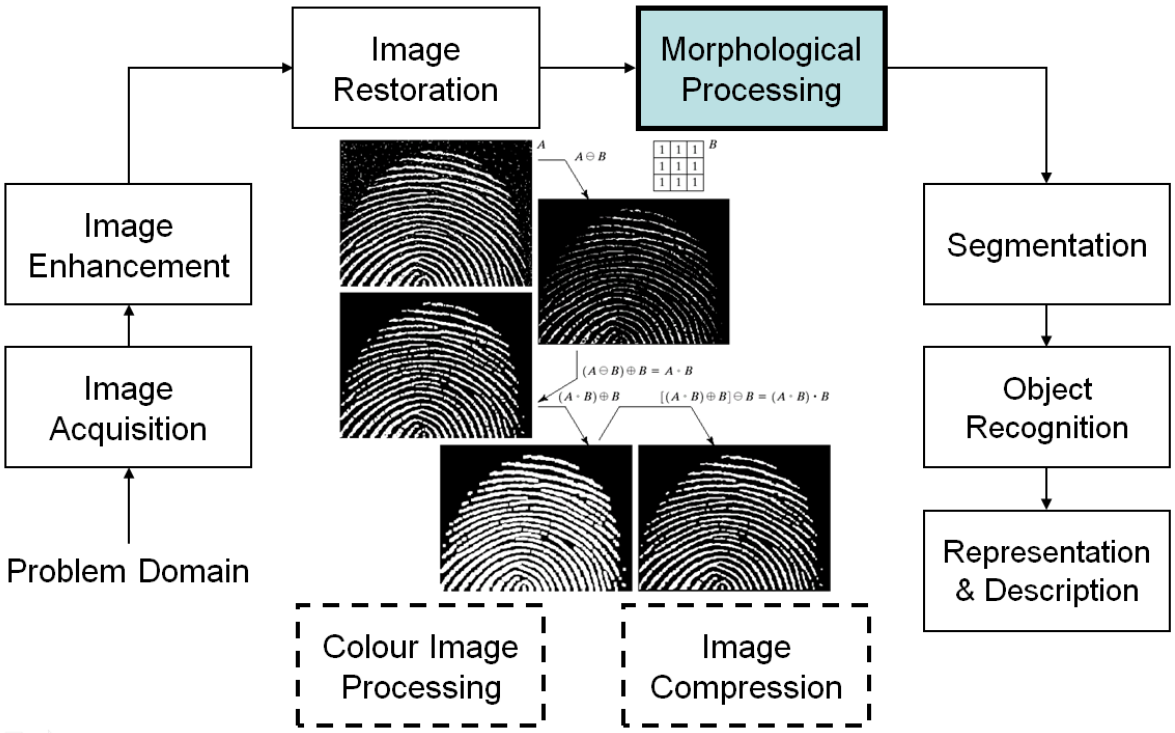
Key Stages in DIP: Image Enhancement



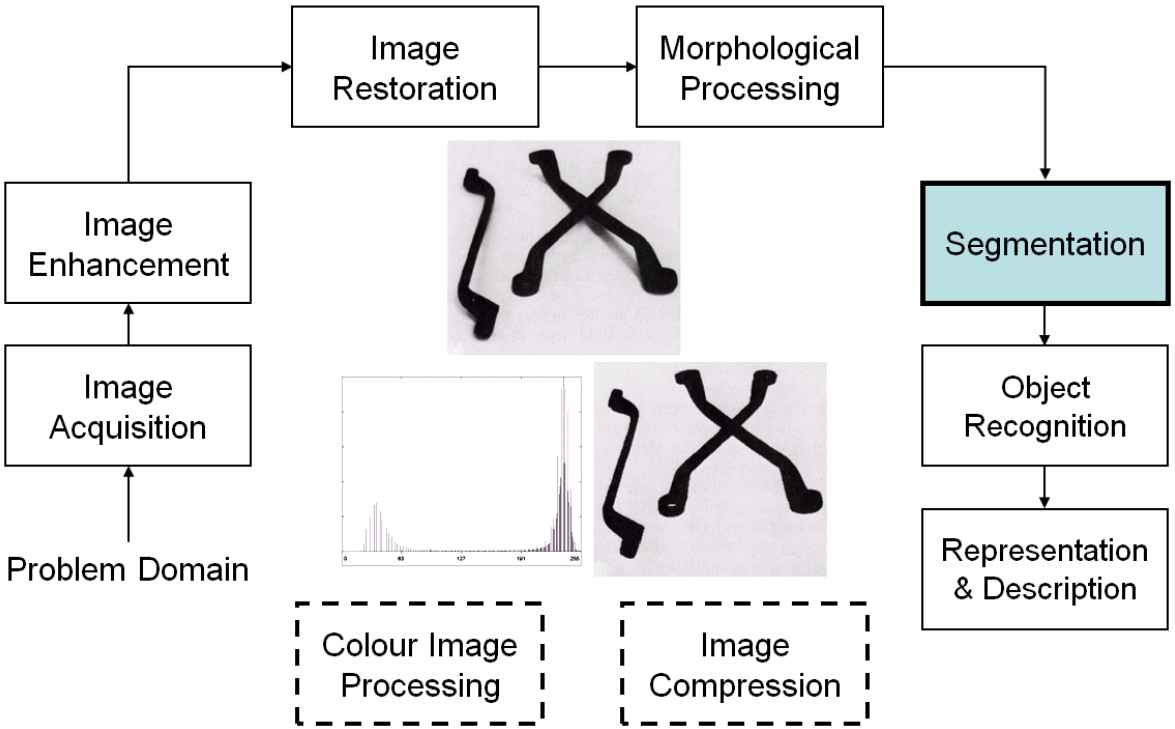
Key Stages in DIP: Image Restoration



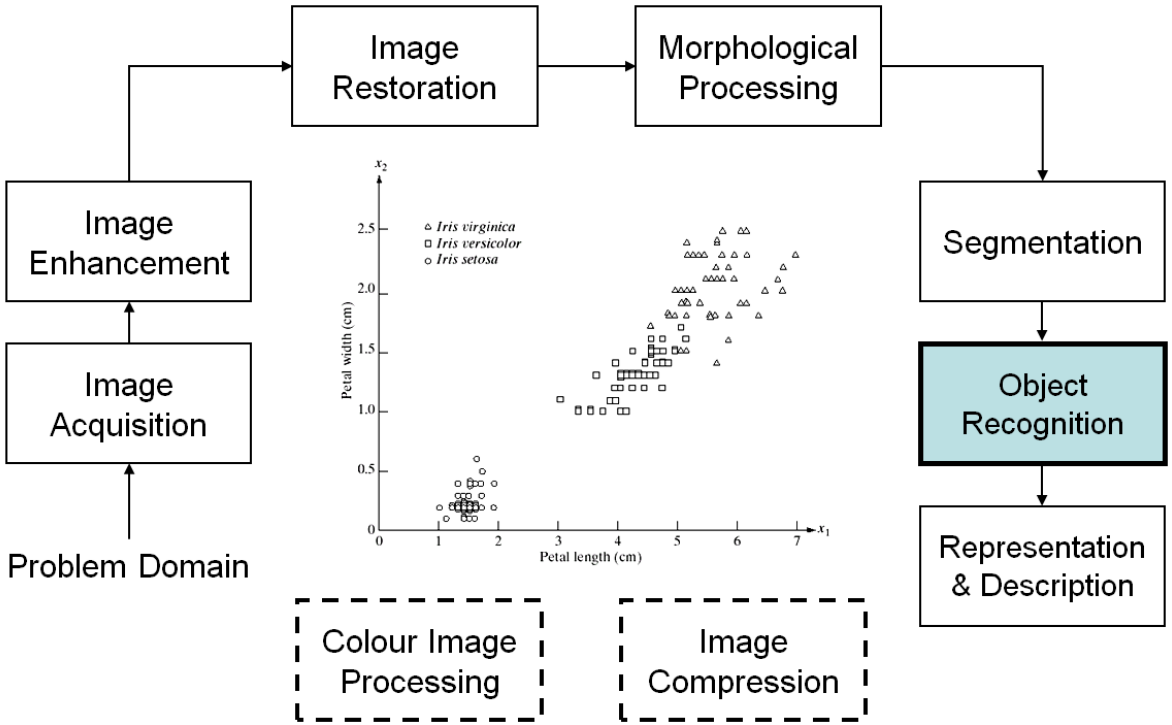
Key Stages in DIP: Morphological Processing



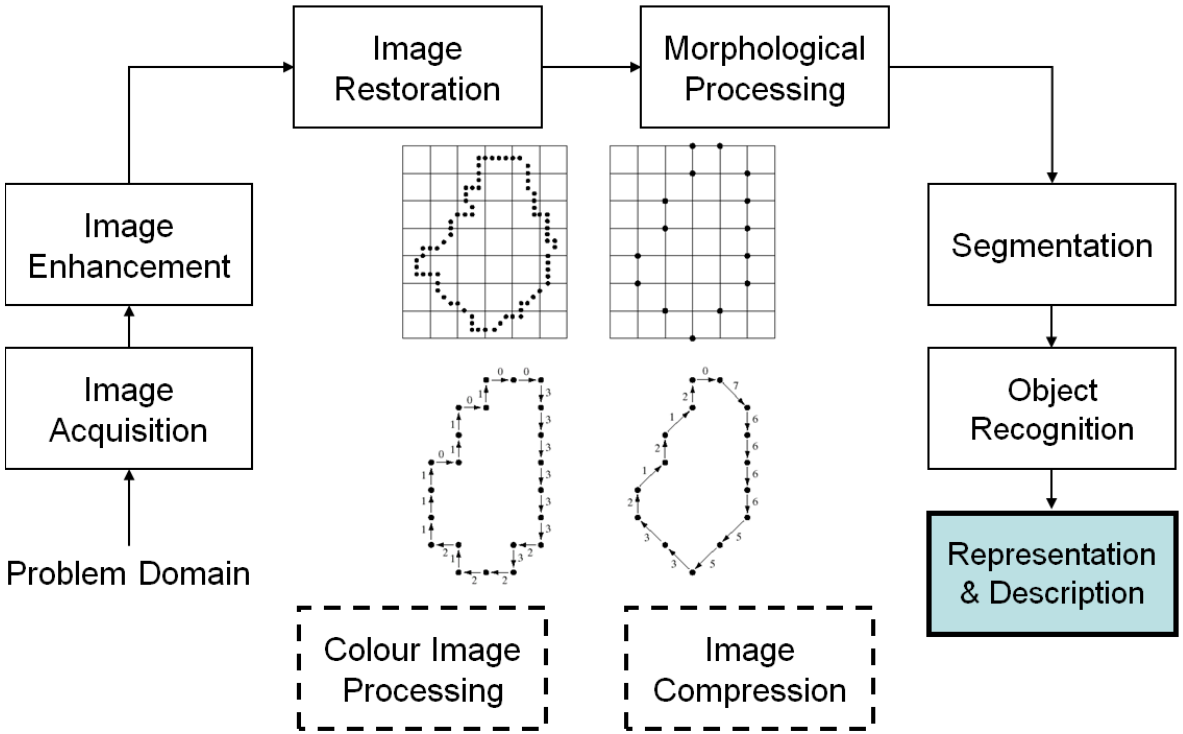
Key Stages in DIP: Image Segmentation



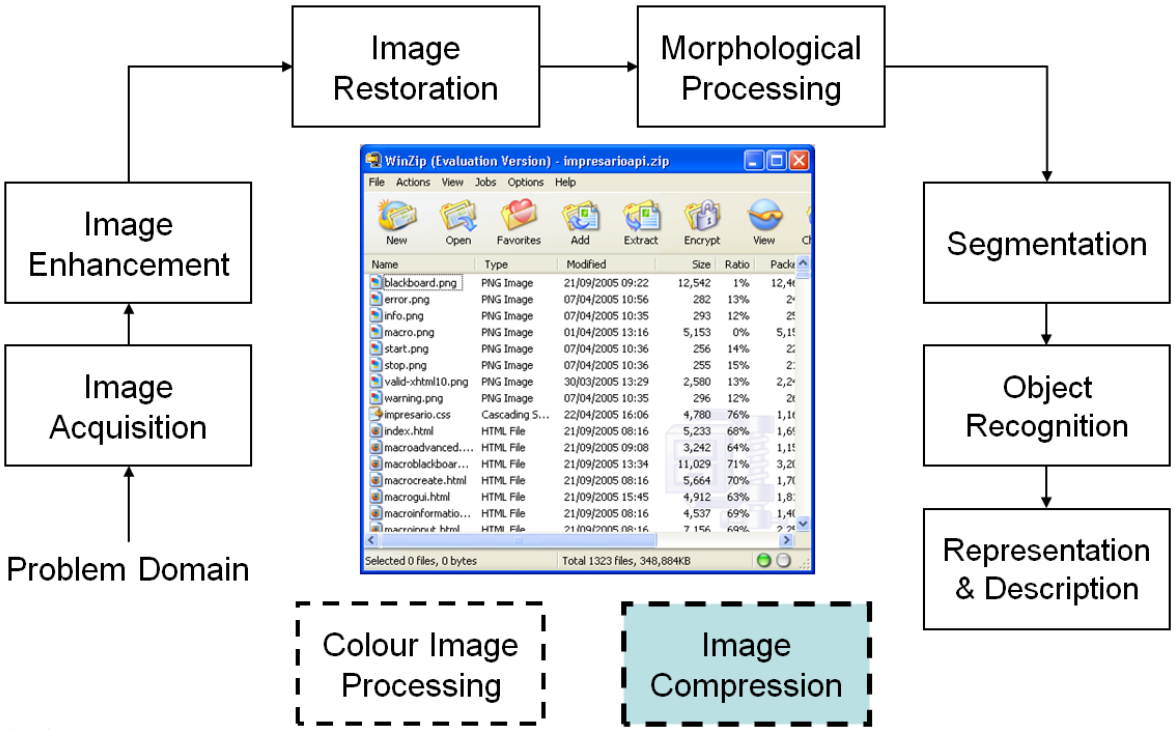
Key Stages in DIP: Object Recognition



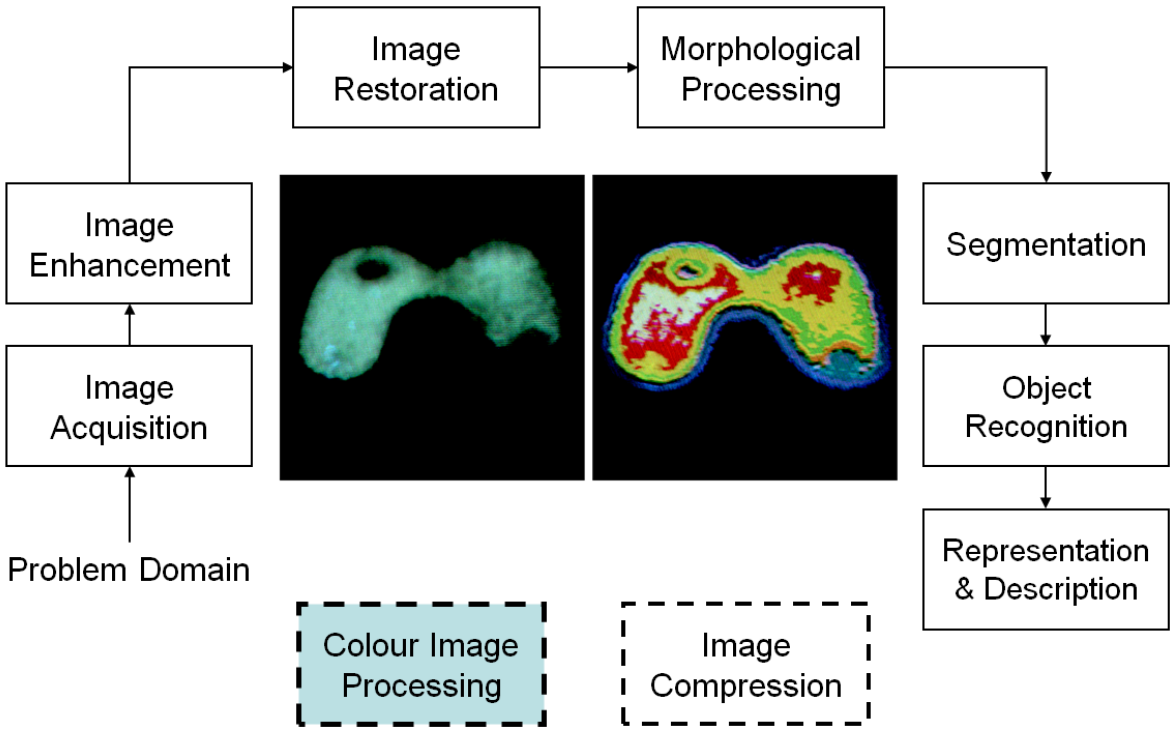
Key Stages in DIP: Representation & Description



Key Stages in DIP: Image compression



Key Stages in DIP: Colour Image Processing



Proposed Projects

The next slides present some projects which you can choose as term projects for this course

UG 1: Human action recognition

- Humans are often the most important, as well as interesting, feature shown in video. This project is concerned with identification of basic human actions such as walking, running, and hand shaking. Refer to the reading list below to find more.
- Readings
  - Action database. KTH.
  - Action recognition datasets. UC Berkeley.

UG 2: Human emotion recognition

- A facial expression results from one or more motions or positions of the muscles of the face. They convey the emotional state of the individual to observers. Facial expressions are a form of nonverbal communication. Paul Ekman suggested six basic emotions related to facial features, namely anger, disgust, fear, joy, sadness, and surprise. The aim of this project is to identify these emotions using facial parts. Speech is a complementary source for this identification.
- Readings
  - Zeng et al (2007). Audio-visual spontaneous emotion recognition. Lecture Notes in Computer Science 4451, Springer.
  - Castellano et al (2008). Emotion recognition through multiple modalities: face, body gesture, speech. Lecture Notes in Computer Science 4868, Springer.
- Datasets
  - HMDB: A Large Human Motion Database, $http://serre-lab.clps.brown.edu/resource/hmdb-a-large-human-motion-database/\#Downloads$
  - List of body parts dataset, $http://liris.cnrs.fr/voir/wiki/doku.php?id=datasets$
  - $ftp://ftp.biosciencedbc.jp/archive/bodyparts3d/$

UG 3: Head gestures identification

- The project aims to identify head nodding and shaking in video streams. They are the most common head gestures to perform conversational functions - head nod for yes, head shake for no, although there seem to exist cultural variations. Detection of these gestures is fundamental to identify human responses visually.
- Readings
  - Lu et al (2005). Head gesture recognition based on Bayesian network LNCS 3522, Springer.
  - Kang et al (2006). Real time head nod and shake detection using HMMs. LNCS 4253, Springer.

UG 4: Gender Classification

- Human's gender can be identified using face information only. Different facial features such as forehead, eyebrows, nose, cheek, top lips length, chin jaw and Adam's apple help in identification of human gender.
- Readings
  - Jian-Gang Wang (2010). Boosting dense SIFT descriptors and shape contexts of face images for gender recognition, IEEE.
  - Duan-Yu Chen; Kuan-Yi Lin (2010). Robust gender recognition for uncontrolled environment of real-life images, IEEE.
  - Hee Lin Wang *et al.*(2010). Real time gender recognition with unaligned face images, IEEE.
  - Ullah, I. *et al.*(2012). Gender recognition from face images with local WLD descriptor, IWSSIP.
  - Min Li (2013). Head-shoulder based gender recognition, ICIP.
  - Santarcangelo, Vito *et al.*(2015). Gender recognition: Methods, datasets and results, ICMEW.
- Datasets
  - Yale Face Dataset, $http://vision.ucsd.edu/content/yale-face-database$
  - Collection of people from Flickr images, $http://chenlab.ece.cornell.edu/people/Andy/ImagesOfGroups.html$
  - The OUI-Adience: Face Image Project, $http://www.openu.ac.il/home/hassner/Adience/data.html$
  - MORPH (Craniofacial Longitudinal Morphological Face Database), $http://www.faceaginggroup.com/morph/$

UG 5: Age Group Identification

- Age can be classified as old, young, child based on face features. The geometrical ratios can be calculated based on the distance, size of certain facial characteristics and estimation of the amount of wrinkles in facial areas.
- Readings
  - Young Ho Kwon *et al.*(1994). Age classification from facial images, IEEE.
  - Haibin Ling *et al.*(2007). A Study of Face Recognition as People Age, ICCV.
  - Bauckhage, C. etal (2010). Age Recognition in the Wild, ICPR.
  - M. Mahdi Dehshibi & Azam Bastanfard (2010). A new algorithm for age recognition from facial images, IEEE.
- Datasets
  - Collection of people from Flickr images, *http : / / chenlab.ece.cornell.edu / people / Andy / ImagesOfGroups.html*
  - The OUI-Adience: Face Image Project, *http : / / www.openu.ac.il / home / hassner / Adience / data.html*
  - MORPH (Craniofacial Longitudinal Morphological Face Database), *http : / / www.faceaginggroup.com / morph /*

UG 6: Body Parts Identification

- Identification of body parts is an important first step for many tasks such as action recognition. A general human body part information is hand, face, head, head, arm, torso and leg.
- Readings
  - Forsyth, D.A. & Fleck, M.M. (1997). Body Plans, IEEE.
  - Dargazany, A. & Nicolescu, M. (2012). Human Body Parts Tracking Using Torso Tracking: Applications to Activity Recognition, ITNG.
  - Wangyang Wei *et al.*(2014). Person Re-identification Based on Human Body Parts Signature, ICDSC.
- Datasets
  - HMDB: A Large Human Motion Database, $http://serre-lab.clps.brown.edu/resource/hmdb-a-large-human-motion-database/\#Downloads$
  - List of body parts dataset, $http://liris.cnrs.fr/voir/wiki/doku.php?id=datasets$
  - $ftp://ftp.biosciencedbc.jp/archive/bodyparts3d/$

UG 7: Identity Recognition

- The aim of this project is to identify the occupation of humans. Social characteristics of human such as social status, connections, and roles in a particular situation draw great attention since they are the essence of social life. Clothing parsing draws attention due to its close relation with people's identity.
- Readings
  - Ming Shao *et al.*(2013). What Do You Do. Occupation Recognition in a Photo via Social Context, ICCV.
  - Z. Song *et al.*(2011). Predicting occupation via human clothing and contexts, ICCV.
- Datasets
  - PubFig database, http://www.cs.columbia.edu/CAVE/databases/pubfig/
  - MPII Human Pose dataset, http://human-pose.mpi-inf.mpg.de/dataset

UG 8: Group/Crowd/People Identification

- A group is defined as two or more people who interact with one another and share similar characteristics. This project aims at identifying multiple humans in a scenes.
- Readings
  - Navneet Dalal & Bill Triggs (2005). Histograms of Oriented Gradients for Human Detection, IEEE.
  - Xi Zhao *et al.*(2009). A People Counting System Based on Face Detection and Tracking in a Video, IEEE.
  - Da-Jinn Wang *et al.*(2009). People Recognition for Entering & Leaving a Video Surveillance Area, ICICIC.
  - Dabrowski, A. *et al.*(2011). People recognition and tracking methods for control of viewpoint in CCTV systems, ECCTD.
  - Brenner, M. & Izquierdo, E. (2013) People recognition in ambiguously labeled Photo Collections, ICME.
- Datasets
  - CBCL pedestrian database, $http://cbcl.mit.edu/software-datasets/PedestrianData.html$
  - Crowd Segmentation Data Set, $http://crcv.ucf.edu/data/crowd.php$
  - Mall Dataset crowd counting dataset, $http://www.eecs.qmul.ac.uk/ccloy/downloads_mall_dataset.html$
  - CUHK Crowd Dataset, $http://www.ee.cuhk.edu.hk/jshao/CUHKcrowd_files/cuhk_crowd_dataset.htm$

UG 9: Dress - Clothing Identification

- Recognition of clothing categories from videos is appealing to emerging applications such as intelligent customer profile analysis and computer aided fashion design. The purpose of this project is to tag the clothing categories such as shirt, jeans, skirt, suit *etc.*from an image.
- Readings
  - Rahman, M.M. *et al.*(2001). Solving a dress problem for a human model recognition, SICE.
  - Sakurai, R. *et al.*(2011). People and clothes recognition based on topic model, SII.
  - Ming Yang, Kai Yu (2011). Real-time clothing recognition in surveillance videos, IEEE.
  - Yamazaki, K *et al.*(2013). A method of state recognition of dressing clothes based on dynamic state matching, SII.
- Datasets
  - Apparel classification with Style, *http* : *//www.vision.ee.ethz.ch/ lbossard/projects/accv*12*/index.html*
  - Fashion 10000 : An Enriched Social Image Dataset for Fashion and Clothing, Data set, *http* : *//traces.cs.umass.edu/index.php/Mmsys/Mmsys*
  - Fashion focused Creative Commons Social database, *http* : *//traces.cs.umass.edu/index.php/Mmsys/Mmsys*
  - Clothing Attributes Dataset, *http* : *//web.stanford.edu/ hchen2/datasets.html#clothingattributedataset*
  - CSIC: Clothing part dataset, *http* : *//www.iri.upc.edu/groups/perception/#clothingDataset*

UG 10: Facial Actions

- The project aims to identify facial actions such as talking, smoking, reading and eating in video streams. Detection of these actions is essential to identify human responses visually.
- Readings
  - Essa, I.A. & Pentland, A. (1994). A vision system for observing and extracting facial action parameters, Computer Vision and Pattern Recognition.
  - Ip, H.H.S. & Chiu, R.C.K. (1994). Evidential reasoning for facial gestures recognition from cartoon images, Intelligent Information Systems.
  - Essa, I.A.; Pentland, A.P. (1997). Coding, analysis, interpretation, and recognition of facial expressions, IEEE.
  - Donato, G. *et al.*(1999). Classifying facial actions, IEEE.
  - Pantic, M. *et al.*(2002). Facial action recognition in face profile image sequences, IEEE.
- Datasets
  - HMDB: A Large Human Motion Database, $http://serre-lab.clps.brown.edu/resource/hmdb-a-large-human-motion-database/\#Downloads$
  - Activity, Event and Action Databases, $http://www.ecse.rpi.edu/homepages/cvrl/database/Activity_Datasets.htm$
  - 60 Facial Recognition Databases, $https://www.kairos.com/blog/60-facial-recognition-databases$

UG 11: Hand Actions

- Hand action recognition pertains to recognizing meaningful expressions of motion by a human, involving the hands. This project aims to identify the action through hands such as holding, carrying, writing, waving, painting and touching.
- Readings
  - Yikai Fang *et al.*(2007). A Real-Time Hand Gesture Recognition Method, IEEE.
  - Dondera, R. *et al.*(2009). Action recognition based on human movement characteristics, WMVC.
  - Liu Yun & Zhang Peng (2009). An Automatic Hand Gesture Recognition System Based on Viola-Jones Method and SVMs, WCSE.
- Datasets
  - Action Database, $http://www.nada.kth.se/cvap/actions/$
  - MPII Human Pose dataset, $http://human-pose.mpi-inf.mpg.de/\#dataset$

UG 12: Body Actions

- Humans are often the most important, as well as interesting, feature shown in video. This project is concerned with identification of basic human actions such as walking, running, and hand shaking.
- Readings
  - Emering, L. *et al.*(1997). Live participant's action recognition for virtual reality interactions, IEEE.
  - Mori, T. *et al.*(2001). Human like action recognition system on whole body motion-captured file, IEEE.
  - Xiaolin Feng & Perona, P., (2002). Human action recognition by sequence of movelet codewords, IEEE.
  - Chippendale, P. (2006). Towards Automatic Body Language Annotation, Automatic Face and Gesture Recognition.
- Datasets
  - HMDB: A Large Human Motion Database, $http://serre-lab.clps.brown.edu/resource/hmdb-a-large-human-motion-database/\#Downloads$
  - Activity, Event and Action Databases, $http://www.ecse.rpi.edu/homepages/cvrl/database/Activity_Datasets.htm$
  - List of dataset, $https://www.cs.utexas.edu/\ chaoyeh/web_action_data/dataset_list.html$
  - Action Database, $http://www.nada.kth.se/cvap/actions/$
  - MPII Human Pose dataset, $http://human-pose.mpi-inf.mpg.de/\#dataset$

UG 13: Man Made Objects

- The aim of this project is to automatically detect and recognize man made objects in a digital image or video. Detection is understood as finding a small rectangular area in the image containing the object. Recognition is the attribution of a class label to the detected object. Examples of man-made objects include cars, house, bike, flag, chair, sofa etc.
- Readings
  - Bhanu, Bir et al (1984). Shape Matching of Two-Dimensional Objects, IEEE.
  - Lin, Wei-Chung, King-Sun Fu (1986). A Syntactic Approach to Three-Dimensional Object Recognition, IEEE.
  - Papageorgiou, C.P. *et al.*(1998). A general framework for object detection, IEEE.
  - Michael Banf & Volker Blanz (2013). Man made structure detection and verification of object recognition in images for the visually impaired, MIRAGE.
- Datasets
  - Zurich building, *http* : $//www.vision.ee.ethz.ch/datasets/index.en.html$
  - CALTECH, *http* : $//www.vision.caltech.edu/html-files/archive.html$
  - *CALTECH*256, *http* : $//www.vision.caltech.edu/Image_Datasets/Caltech256/$
  - *CALTECH*101, *http* : $//www.vision.caltech.edu/Image_Datasets/Caltech101/\#Download$
  - *MIT* − *AdobeFiveKDataset*, *http* : $//groups.csail.mit.edu/graphics/fivek_dataset/$

UG 14: Natural Objects

- The intend of this project is to automatically detect and recognize natural objects in a digital image or video. Natural objects include tree, plants, flowers, grass, sun, dog, insects etc.
- Readings
  - Efenberger, W. & Graefe, V. (1996). Distance-invariant object recognition in natural scenes, IROS.
  - Yongmei Cheng *et al.*(2009). Natural Object Recognition Using the Combination of Gaussian Model and Region Growing, CISP.
  - Wang *et al.*(2009). Learning Models for Object Recognition from Natural Language Descriptions, *http* : //*www*.*comp*.*leeds*.*ac*.*uk*/*me*/*Publications*/*bmvc*09.*pdf*
- Datasets
  - CALTECH (birds), *http* : //*www*.*vision*.*caltech*.*edu*/*visipedia*/*CUB* − 200 − 2011.*html*
  - MIT-Adobe FiveK Dataset, *http* : //*groups*.*csail*.*mit*.*edu*/*graphics*/*fivek*$_d$*ataset*/
  - PASCAL Context Dataset, *http* : //*www*.*stat*.*ucla*.*edu*/ *ccvl*/*datasets*/*pascal* − *context*/

UG 15: Scene Settings

- Classifying an image into a particular scene setting is computationally a very challenging task. It is difficult due to vast range of variations in these scene categories. An outdoor image can be of beach, urban scene, houses, road traffic etc and an indoor image can have scene from office, bedrooms, bakery, hotels.
- Readings
  - Jiming Li & Yuntao Qian (2008). Automatic scene recognition for digital camera by semantic features, ICWAPR.
  - Ramasamy, T. *et al.*(2011). Detecting background setting for dynamic scene, CSPA.
  - Sangseung Kang *et al.*(2011). Indoor location recognition system using environmental sensors, ICCAS.
- Datasets
  - Outdoor Scene Attribute (SceneAtt) dataset, $http://www.stat.ucla.edu/\ shuo.wang/SWang_Dataset.html$
  - SUN Attribute Database, $http://cs.brown.edu/\ gen/sunattributes.html$
  - Imperial College London: $Livingroomdataset$, $http://www.doc.ic.ac.uk/\ ahanda/VaFRIC/living_room.html$
  - CALTECH, $http://www.vision.caltech.edu/html-files/archive.html$
  - MIT-Adobe FiveK Dataset, $http://groups.csail.mit.edu/graphics/fivek_dataset/$

UG 16: Location Identification

- Place and location information (*e.g.*, room, office, hospital, cafeteria, road, street, valley) are important as they show the position of humans or other objects in the scene.
- Readings
  - Schindler, G. *et al.*(2007). City-Scale Location Recognition, CVPR.
  - Feryanto, A. & Supriana, I. (2011). Location recognition using detected objects in an image, ICEEI.
- Datasets
  - The Oxford Building Dataset, *http* : $//www.robots.ox.ac.uk/\ vgg/data/oxbuildings/$
  - SUN Database, *http* : $//groups.csail.mit.edu/vision/SUN/$
  - PASCAL-Context Dataset, *http* : $//www.stat.ucla.edu/\ ccvl/datasets/pascal-context/$

Proposed Projects

- You can propose your own projects.
- Projects are assigned on first come first served basis.
- There should be no repetitions for the projects.