INDEX

1. Various fields using machine learning
2. History of computer vision
    1. Evolutions big bang
    2. Camera obscura
    3. Study of vision
    4. Block world
    5. Summer computer vision project
    6. Generalized Cylinder structure VS Picture structure
    7. David Lowe
    8. Normalized cut
    9. Face detection – Cieder boost algorithm
    10. Fuji Film
    11. SIFT and Object detection
    12. Spatial pyramid matching
    13. SVM
    14. Histogram of gradients
    15. Deformable part model
    16. PASCAL Vision Object challenge
    17. ImageNet

VARIOUS FIELDS USING MACHINE LEARNING

- Neuroscience
- Cognitive sciences
- Graphics
- Information retrieval
- Machine learning
- Robotics
- Natural language processing
- Optics

HISTORY OF COMPUTER VISION

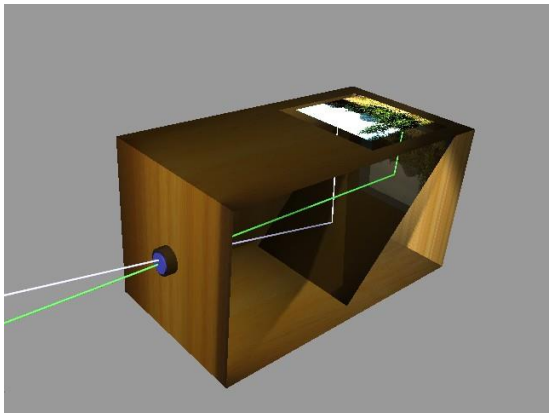
2.1 Evolutions big bang

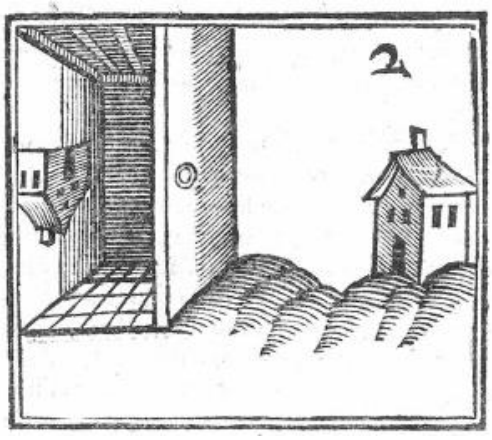
Year: 543 million years agoScientist: Andre parker

The sudden appearance of range of animals during big bang evolution was due to the development of vision. Due to vision led to development of hard body parts.

2.2 Camera obscura

Natural phenomena occurs when an image in a screen was projected by





inverted and reversed on other end through a small hole in screen.

Scientist history

- Gemma frisius – 1545
    - Is a mathematician who created globes, applied his improved mathematics in surveying and navigation
- Leonardo da Vinci – 16th Century
    - He is an Italian interested in invention, drawing, painting, sculpture, architecture, science, music, mathematics, engineering, literature, anatomy, geology, astronomy, botany, paleontology and cartography. He was greatest painter.

2.3 Study of vision

Year: 1959

Search value: Study of vision Hubel and Wiesel

Link: https://www.youtube.com/watch?v=Cw5PKV9Rj3o

https://www.youtube.com/watch?v=IOHayh06LJ4

Authors: David Hubel and Torstein Wiesel

They study based on cat retina, as they moved bright line across retina.

They noticed that

1. Neurons fires only when line was perpendicular to retina.
2. The activity of these neurons changed depending on the orientation of the line.
3. Sometimes neurons fires only when line was moving in a particular direction.

Simple cells – Response to light orientation

Complex cells - Response to light orientation and other

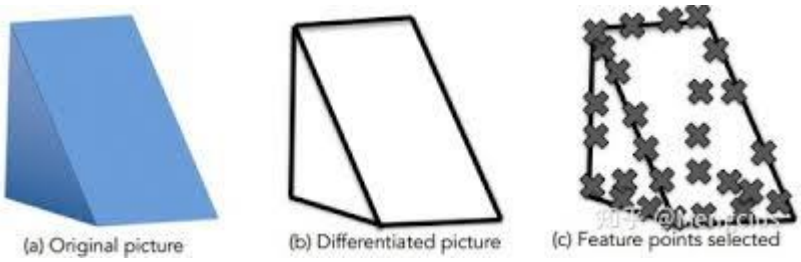
Hyper complex cells - Response to movement with an end point

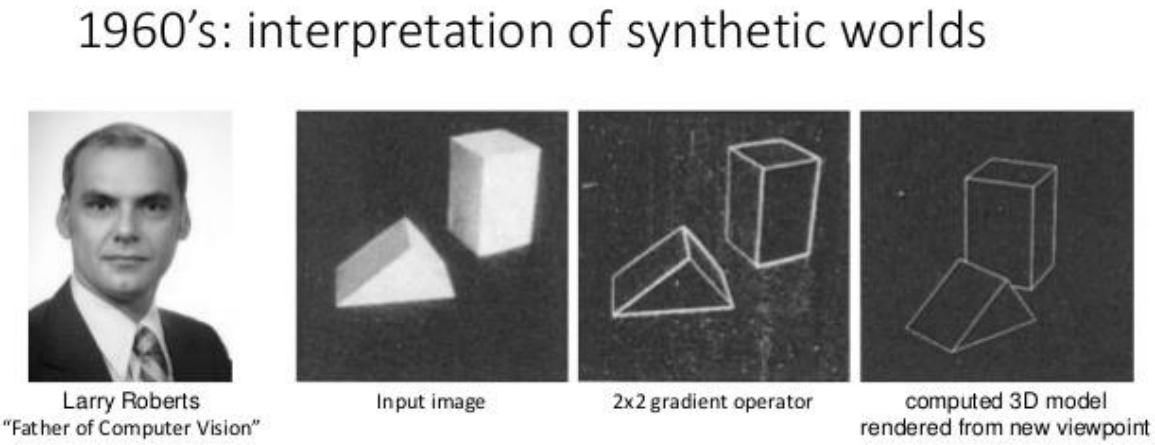
2.4 Block world

Year: 1963

Author: Larry Roberts

Considered as first PhD in computer vision algorithm where visual world simplified into Geometrical space. Tried to make 3D geometry by arrangements of blocks





2.5 Summer vision Project

Year: 1966

Author: MIT

The goal was to summer workers effectively in the construction of significant part of a visual system.

Field: Pattern recognition.

David Marr (MIT – Computer vision scientist)

Steps

1. Input image
2. Edge Images
    a. Blobs, edges, ends, lines etc.
3. 2 1/2 – D Sketch
    a. Place all together to form a surface, discontinued.

2.6 Generalized Cylinder structure VS Picture structure

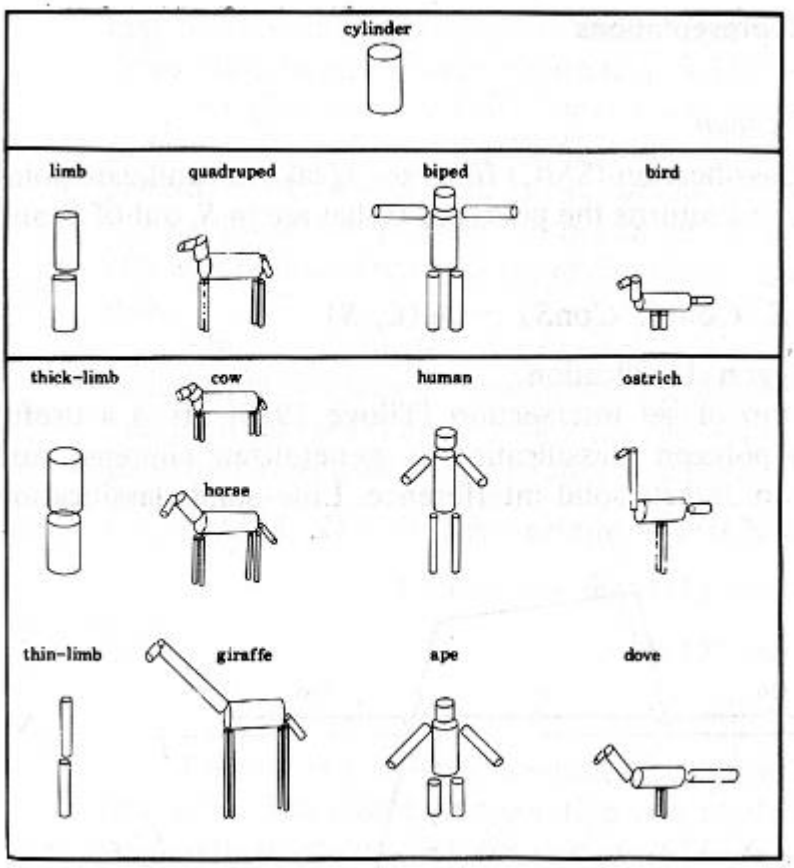
Generalized structure

Year: 1979

Author: Brooks and Binford

Link: http://homepages.inf.ed.ac.uk/rbf/BOOKS/BANDB/LIB/bandb9_34.pdf

In generalized cylinder approach, a human can be represented in geometrical view by cylindrical shape. Diagram given below,



Picture structure

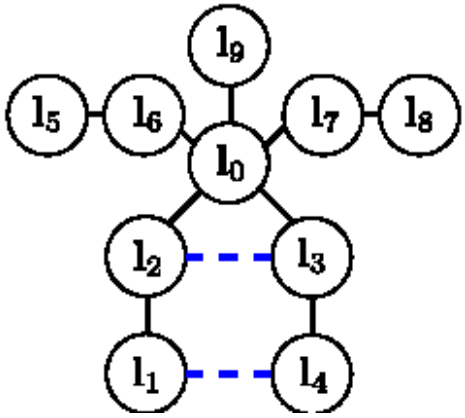
Author: Fischler and Elschlager

Reference: http://www.cs.cornell.edu/~dph/papers/pict-struct-ijcv.pdf

In picture representation, a human can be represented in geometrical view by critical parts in the object and distance between critical parts and its shape.

An object is modeled by collection of parts arranged in deformable configuration. Each part encodes visual properties of the object. Deformable configuration is characterized by spring like connection between certain pairs of parts. The best match found by minimizing an energy function that measures both a match cost for each part and a deformation cost for each pair of connected parts.

5 Visualization of the
ors between body parts in
nodel. *Solid lines*
espond kinematic factors
dashed lines correspond to
lsive factor



2.7 David Lowe

He is Canadian computer scientist, he tells that we can recognize objects by constructing lines, edges and its combinations.

2.8 Normalized cut

Year: 1997

Authors: Jitendra malik and Jianbo shi.

Reference: http://www.cis.upenn.edu/~jshi/papers/pami_ncut.pdf

Extract image impression of an image. Treat image segmentation as a graph partitioning problem and proposed a novel globe criteria.

Keywords: Grouping, Image segmentation, graph partitioning.

The normalized cut measures the dissimilarity and similarity between within the groups.

2.9 Face detection – Cider boost algorithm

Authors: Viola and Jones

References: https://www.cs.cmu.edu/~efros/courses/LBMV07/Papers/viola-cvpr-01.pdf

There are three approaches mainly using,

- Integral image
    - Image representations, which can rapidly compute features of image

    - Doing some operations on each pixel of image
- Learning algorithm
    - Based on adaBoost, which selects small number of visual features from larger sets yields extremely efficient algorithms.
    - Classifier for constructing important features using adaBoost.
- Cascade
    - Remove background parts of images quickly, while spending more computation on object like regions.
    - Combining more complex classifiers in a cascade by focusing on important regions of image

2.10 Fuji Film

Year: 2006

Name: Fujifilm FinePix S6500fd

Digital camera from FUJI, they comes with first time with face detection feature.

2.11 SIFT and object detection

Year: 1999

Author: David Lowe

Reference: https://www.cs.ubc.ca/~lowe/papers/iccv99.pdf

https://www.vocal.com/video/scale-invariant-feature-transform-sift-for-object-detection/

Matching entire object between two step sizes.

SIFT (Scale invariant feature transform) is feature extraction tool which are stable for rotation, translation and scaling and somewhat invariant changes in illumination and camera view point.

2.12 Spatial pyramid matching

Year: 2006

Author: Svetlana Lazebnik, Jean Ponce, Cordelia Schmid

Reference:
https://www.cs.unc.edu/~lazebnik/publications/pyramid_chapter.pdf

This technique partition image into fine sub-regions and computing histograms of local features found on each sub region. The resulting spatial pyramid consist of bag-of-features which enough to scene categorization tasks.

2.13 SVM

Year: Introduced in 1992

For details. Search Google.

2.14 Histogram of Oriented Gradients (HoG)

Year: 2005

Authors: Dalal and Triggs

Reference: https://en.wikipedia.org/wiki/Histogram_of_oriented_gradients

HOG is used technique to count occurrence of gradient orientations in localized portion of image.

A Gradient is a vector, slope is scalar. Gradients are meaningful in multivariable functions, where gradient is a vector of partial derivatives. In one dimensional, they both look same.

2.15 Deformable part model

Year: 2009

Authors: Felzanswalb, Ramanan

Reference:
http://vision.stanford.edu/teaching/cs231b_spring1213/slides/dpm-slides-ross-girshick.pdf

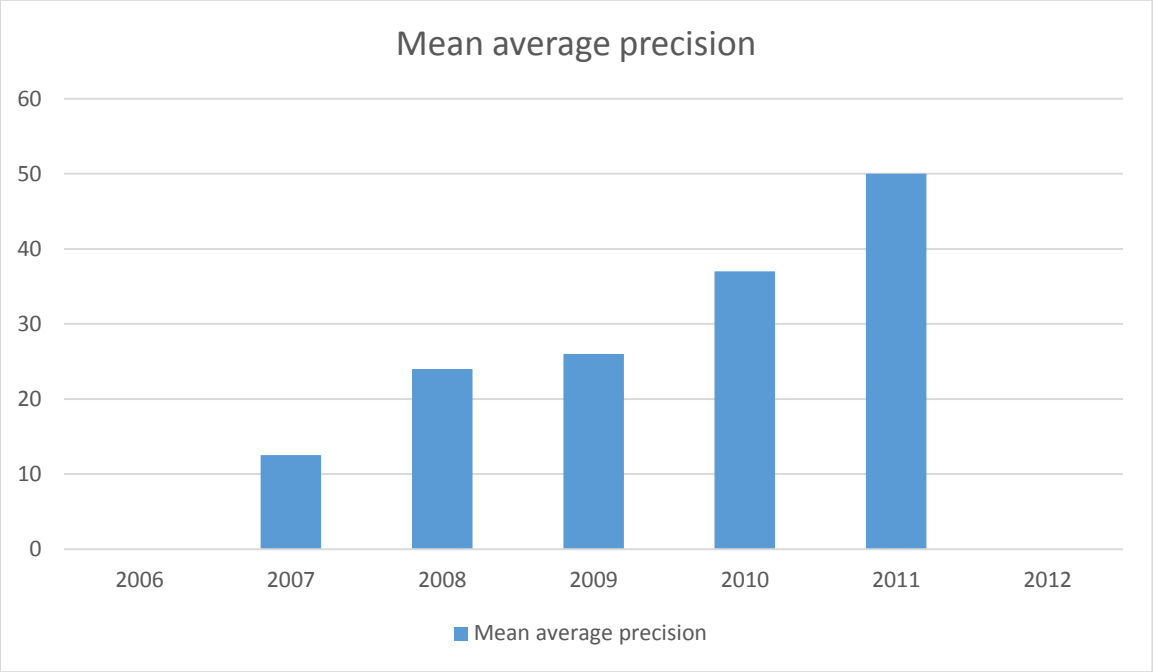
2.16 PASCAL Object challenge

Everingham et al – 2006 - 2012

20 object categories

Object recognition

1000 images/category



2.17 ImageNet

Year: 2009

Authors: Deng Dong, Socher, Li Li, Fei Fei

22K categories

14M images

International challenge conducting ever year

WorldNet – large set of object classes – It is a dictionary of ImageNet organized in well manner.

ILSVRS(ImageNet large scale visual recognition challenge)