Ignacio Devega

Assignment A01

Professor Anna Devarakonda

January 18, 2025

Historical Timeline of Computer Vision

|  |  |
| --- | --- |
| 1700 to 1900 | Scientific community were understanding the principles of light and vision until the first camera system was created by Kodak in 1884. |
| 1957 | The first digital image scanner was developed by Dr. Russell A. Kirsch. |
| 1962 | The research of David Hubel and Torsten Wiesel on Visual Cortex has an impact on the development of computer vision algorithms. |
| 1963 | Larry Roberts, the father of Computer Vision, presented his thesis on machine perception of three-dimensional solids. |
| 1966 | The first summer vision project was done. The purpose was to identify objects, background and chaos in the picture. |
| 1967 | Woodrow W. Bledsoe and I. Kanter presented their thesis on the secret history of facial recognition |
| 1968 | Ivan Sutherland, the computer scientist, developed the Sword of Damocles, a head-mounted display system. |
| 1972 | Richard Duda and Peter Hart developed the Hough Transform. |
| 1973 | Michael A. Fischler and Robert A. Elschlager introduced the concept of “pictorial structure”. |
| 1980 | Automatix, Inc was the pioneer in computer vision and industrial automation. They specialized in developing computer vision and robotic system into the manufacturing process. |
| 1980 | Kunihiko Fukushima developed the NeoCognitron, a neural network model for pattern recognition and image processing. |
| 1981 | Martin A. Fischler and Robert C. Bolles introduced the RanSAC algorithm in their paper “Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography.” |
| 1981 | Bruce D. Lucas and Takeo Kanade introduced the Lucas-Kanade Optical Flow argorithm to estimate motion of objects. |
| 1983 | The book “Vision’ was published by David Marr. It has created a great impact on computer vision and visual perception. |
| 1986 | “The variation approach to shape from shading” was published by Horn and Brooks for 3D shape recovery from images. |
| 1987 | The Deriche edge detector was another addition to Computer Vision. It is a recursive filter to suppress noise in the image. |
| 1988 | Active contour Model (ACM) is used for image segmentation and boundary detection. It’s often called “Snakes”. It was introduced by Michael Kass, Andrew Witkin, and Demetri Terzopoulos. |
| 1989 | Mumford-Shah functional is a mathematical model used for image segmentation and restoration. |
| 1991 | Matthew Turk and Alex Pentland developed the Eigenfaces algorithm. It is a facial recognition method based on principal component analysis (PCA). |
| 1993 | Tony Lindeberg introduced the Scale-space blob detection. It is essential for object recognition and shape analysis. |
| 1999 | Michael Carson introduced “Blobworld: A System for Region-based Image Indexing and Retrieval.” |
| 1999 | SIFT (Scale-Invariant Feature Transform) is a valuable tool for object recognition and image stitching. |
| 2000 | OpenCV (Open-Source Computer Vision Library) was developed by Intel. It’s an open-source computer vision and machine learning software library. It is suitable for video analysis, robotics, and augmented reality. |
| 2001 | Haar Cascades are a machine learning object detection method used to identify objects in images or video. They were developed by Paul Viola and Michael Jones. |
| 2004 | The Viola Jones Face Detection Model which was developed by Paul Viola and Michael Jones, is a pioneering and influential approach for detecting faces in images and videos. |
| 2005 | HOG (Histogram of Oriented Gradients) was first introduced by Robert K. McConnel in 1986. Later HOG algorithm was introduced by Navneet Dalal and Bill Triggs. |
| 2005 | Visual SLAM (Simultaneous Localization and Mapping) is a subset of SLAM. The algorithm is done by matching observed features with the features in the map. |
| 2006 | SURF (Speeded-Up Robust Features) was developed by Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. It is a feature detection and description algorithm in computer vision. |
| 2007 | PCVR (Principal Curvature-Based Region) is a computer vision and image processing technique. It is effective in object detection, image segmentation, and medical image analysis. |
| 2009 | The PASCAL VOC (Visual Object Classes) and ImageNet competitions provide large-scale datasets and benchmark challenges. These competitions played a crucial role in the advancements in computer vision algorithm. |
| 2010 | ILSVRC (ImageNet Large Scale Visual Recognition Challenge) is an annual competition in computer vision and image recognition. Participants had to develop algorithms and models to correctly classify objects within images. |
| 2010 | Microsoft released the Kinetic sensor. It’s a combination of RGB (color) and depth information for real time 3D object recognition and tracking. |
| 2012 | Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton developed the AlexNet. It was the winning entry in the ImageNet Large Scale Visual Recognition Challenge in 2012. |
| 2014 | COCO Data Set by Microsoft has become a prominent dataset in Computer Vision. It has diversity of object categories and complex scenes. |
| 2014 | GANs (Generative Adversarial Network) was introduced by Ian Goodfellow and his colleagues. It’s a deep learning framework. It has become a powerful tool for generating data. |
| 2014 | Another tool that materialized was the VGG. It is one of the classic deep learning architectures widely used because of its simplicity and effectiveness. |
| 2015 | SMPL (Skinned Multi-Person Linear) is primarily used for modeling and animating the human body in a realistic and efficient manner. |
| 2015 | Alexander Mordvintsev , Google engineer, created DeepDream. It’s a convolutional neural networkmethod to enhance and modify images in a unique and psychedelic way. |
| 2015 | Neural Style Transfer is a deep learning technique that combines the content of one image with the artistic style of another. |
| 2015 | Shaoqing Ren, Kaiming Hi, Ross Girshick, and Jian Sun introduced the Faster R-CNN. It is a deep learning-based object detection framework. |
| 2017 | OpenPose was developed by the CMU Perceptual Computing Lab. It is used to identify and locate key body parts and joints in images or video frames. |
| 2017 | Kaiming He, Georgia Gkioxari, Piotr Dollar, and Ross Girshick developed the Mask R-CNN. It is an extension of the Faster R-CNN object detection framework. It’s a powerful tool for various computer vision tasks that require detailed object segmentation. |
| 2018 | YOLO V3 offered real-time object detection capabilities. It has improved efficiency and speed compared to traditional two-stage object detection systems. |
| 2019 | Google’s EfficientNet is the more recent architecture designed to balance between model accuracy and computational efficiency. It is often used when computational resources are limited. |
| 2020 | VIT (Vision Transformers) are a class of deep learning models that apply the Transformer architecture. |
| 2020 | OpenAI developed the GP-3. It is a large language model with natural language processing capabilities. It’s designed for text-related tasks and vision language tasks. |

Reference:

https://letsdatascience.com/learn/history/history-of-computer-vision/