मोतीलाल नेहरू राष्ट्रीय प्रौद्योगिकी संस्थान इलाहाबाद

Motilal Nehru National Institute of Technology Allahabad

An Institute of National Importance as Declared by NIT Act, GOI, 2007

COMPUTER AND MACHINE VISION

Dr. Dushyant Kumar Singh

Assistant Professor

Department of Computer Science & Engineering

dushyant@mnnit.ac.in, +91-9359133388

http://mnnit.ac.in/images/csedfp/dushantcsed/

OUTLINES

- Introduction
- Computer Vision Applications
- Object Detection
- Object Detection Approaches
  - Viola-Jones Approach
  - Frame-Differencing
  - Skin Color Modeling
  - Optical Flow approach
- Experimental Results
- References

INTRODUCTION

- Computer/Machine Vision is the ability of a computer or an automated machine to see.
- It is concerned with the theory for building artificial systems that obtains information from images.
- It is the technology to replace or complement manual inspections and measurements with digital cameras and image processing.

The word Computer Vision is alternatively used for Machine Vision

Introduction

- Vision make computers/machines understand images and videos.

What kind of scene?

Where are the cars?

How far is the building?

…

INTRODUCTION

- Computer vision Vs human vision

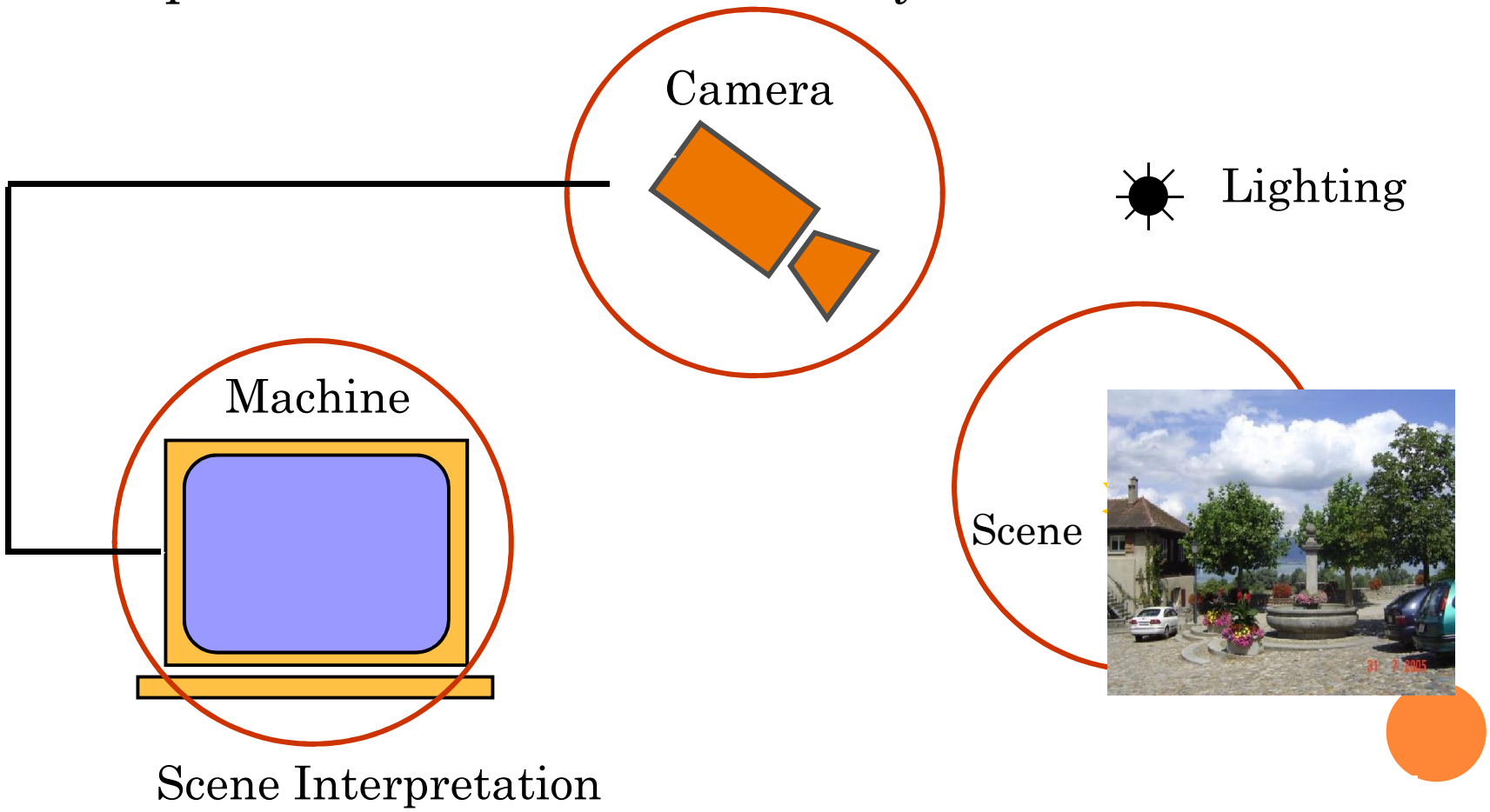
What we see

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

What a computer sees

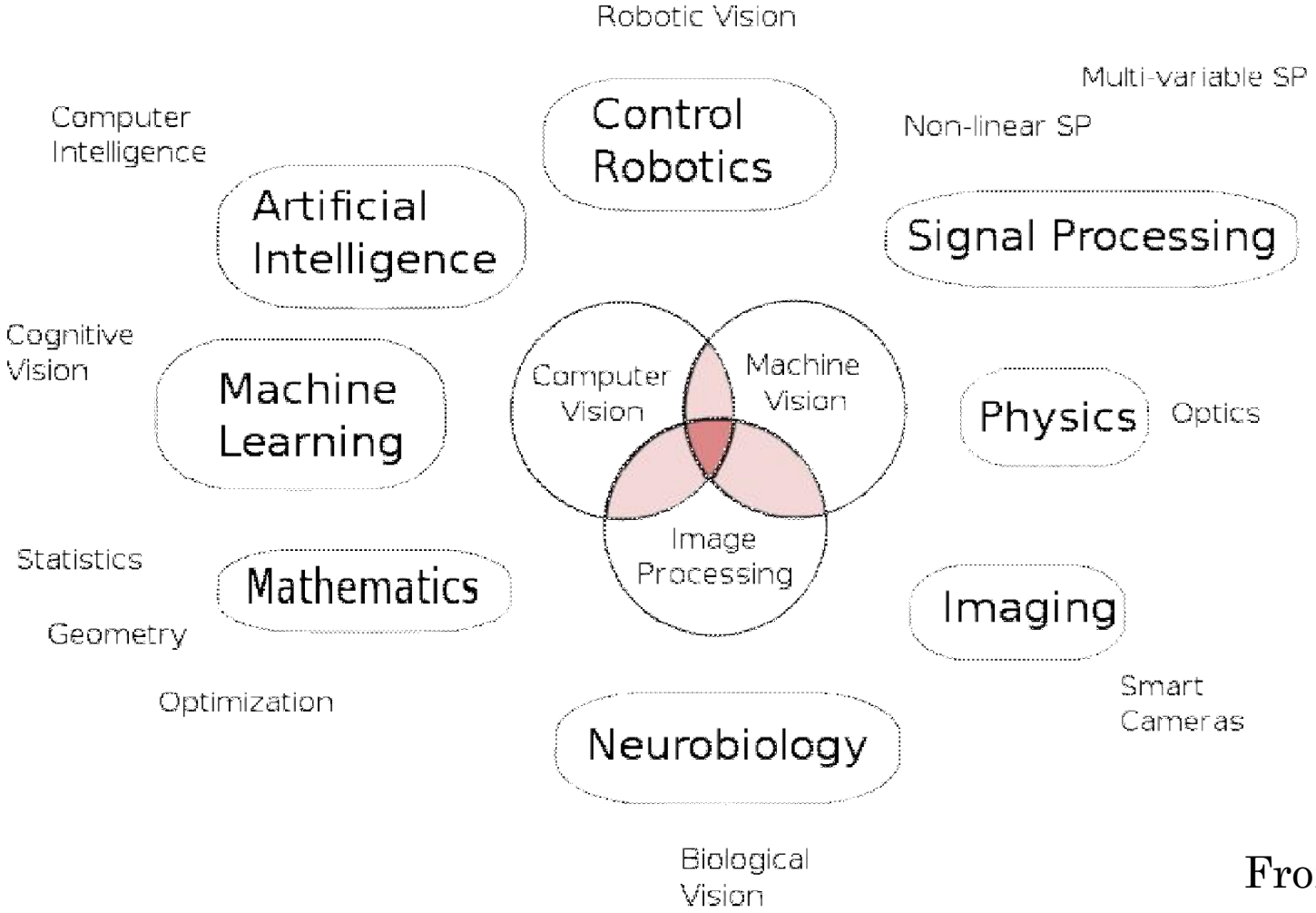
Introduction

- Components of Machine Vision System



Introduction

- Vision is multidisciplinary

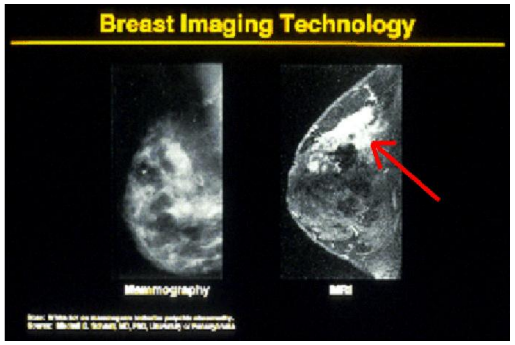


From wiki

Introduction

- Why machine vision matters

Safety

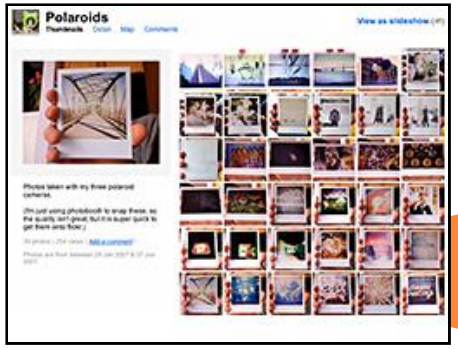


Health

Security

Comfort

Fun



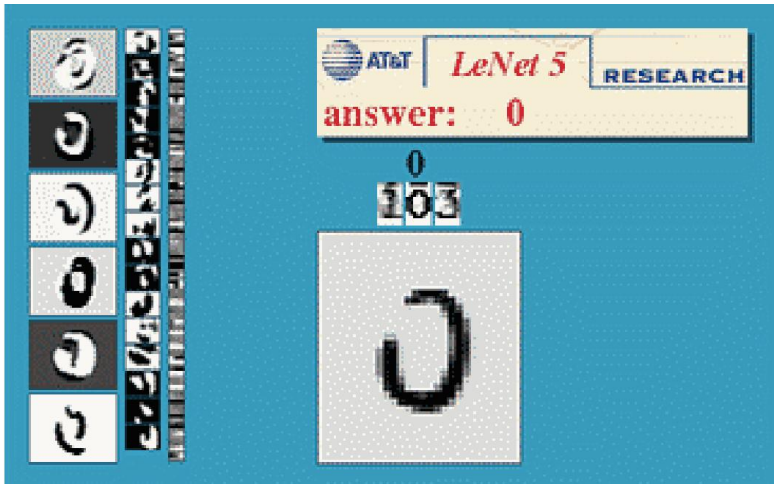
Access

Introduction

- How vision is used
- Examples of state-of-the-art

Optical character recognition (OCR)

Technology to convert scanned docs to text



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

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

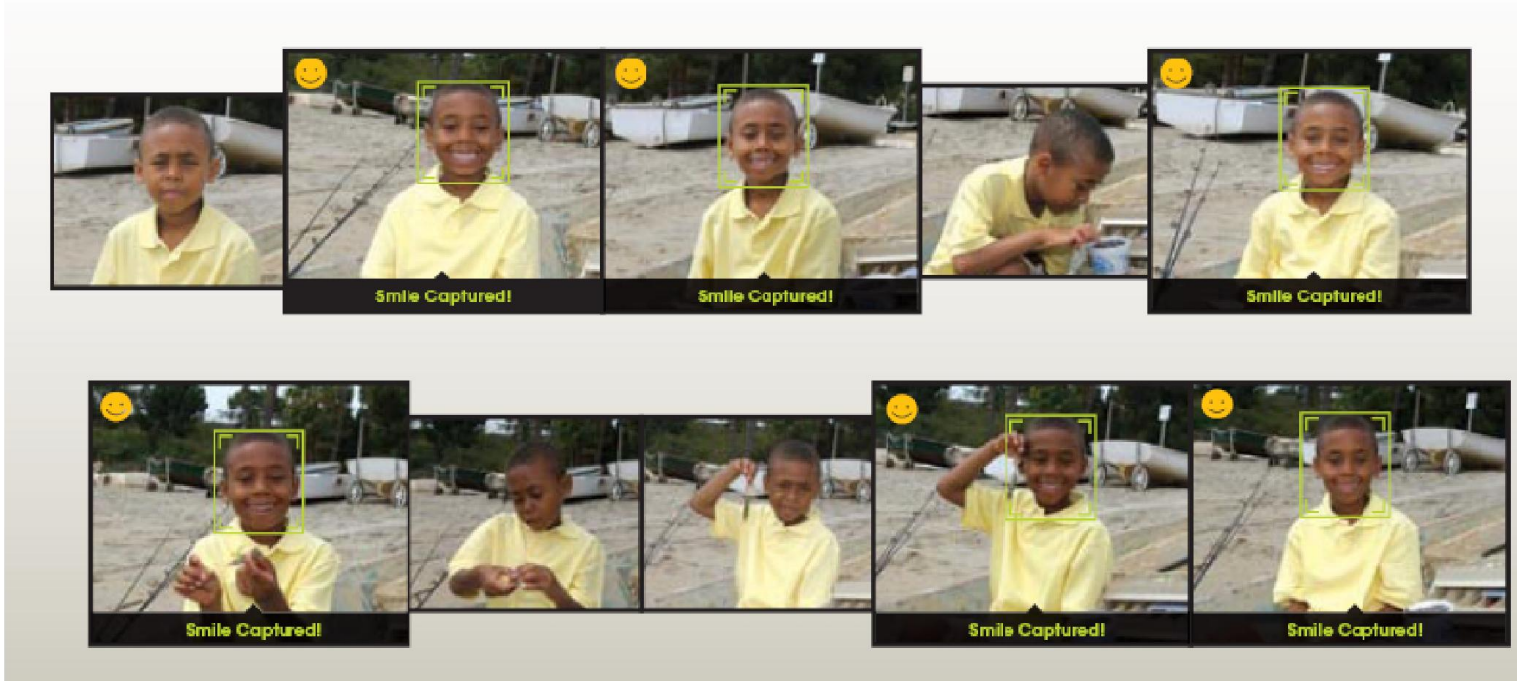
Face detection

- Many new digital cameras now detect faces
  - Canon, Sony, Fuji, …

SMILE DETECTION

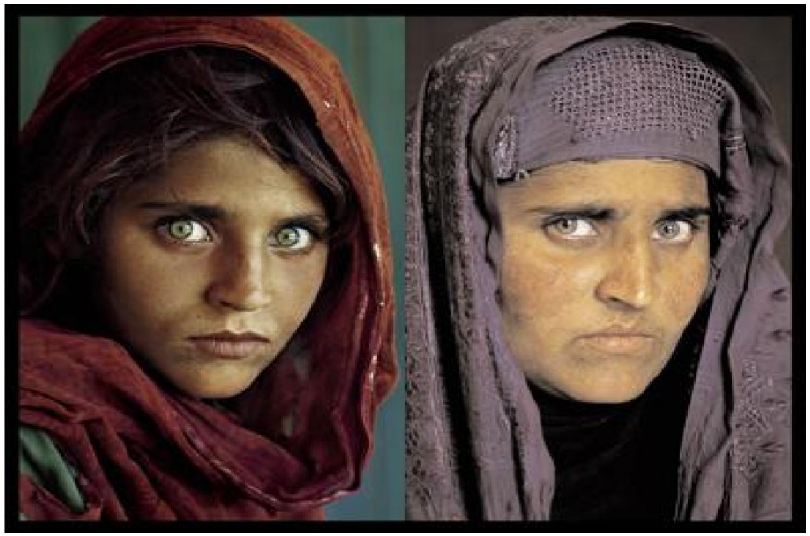
The Smile Shutter flow

Imagine a camera smart enough to catch every smile! In Smile Shutter Mode, your Cyber-shot® camera can automatically trip the shutter at just the right instant to catch the perfect expression.

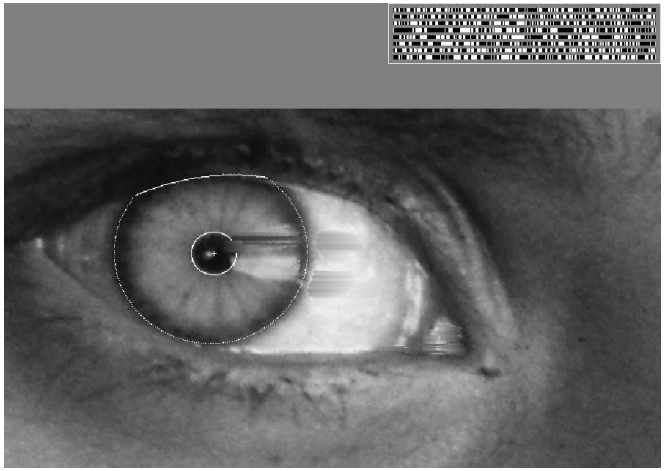


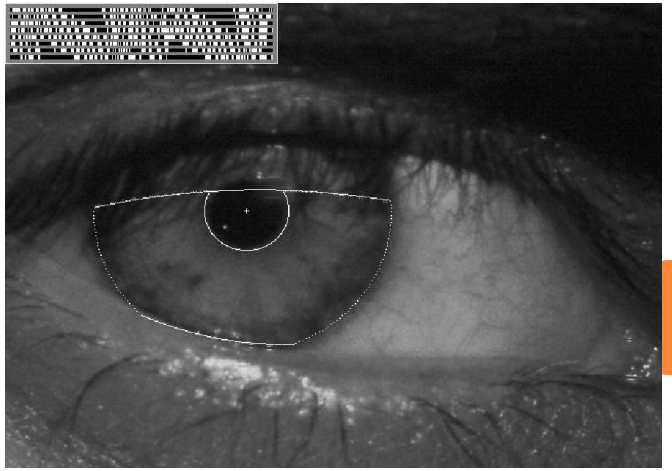
Sony Cyber-shot® T70 Digital Still Camera

VISION-BASED BIOMETRICS



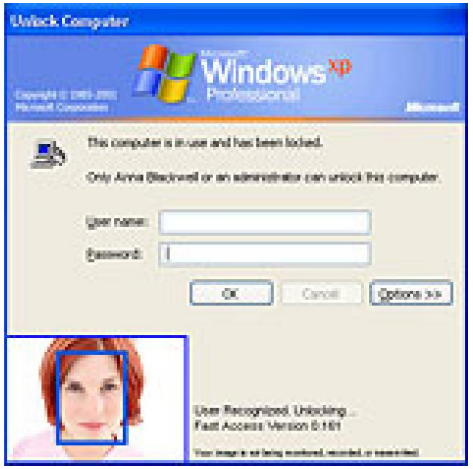
"*How the Afghan Girl was Identified by Her Iris Patterns*" Read the [story](#) [wikipedia](#)





Login without a password…

Fingerprint scanners on many new laptops, other devices



Face recognition systems now beginning to appear more widely
http://www.sensiblevision.com/

Object recognition (in mobile phones)

Point & Find, Nokia

Google Goggles

SPECIAL EFFECTS: SHAPE CAPTURE

The Matrix movies, ESC Entertainment, XYZRGB, NRC

Sports

Sportvision first down line
Nice explanation on www.howstuffworks.com

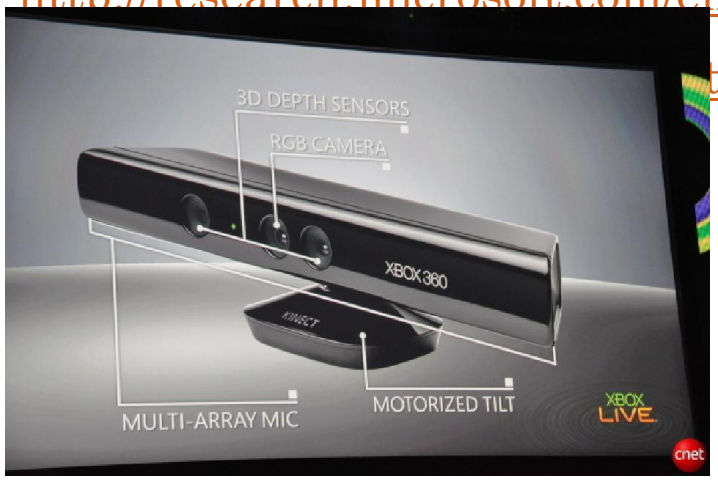
http://www.sportvision.com/video.html

Google cars

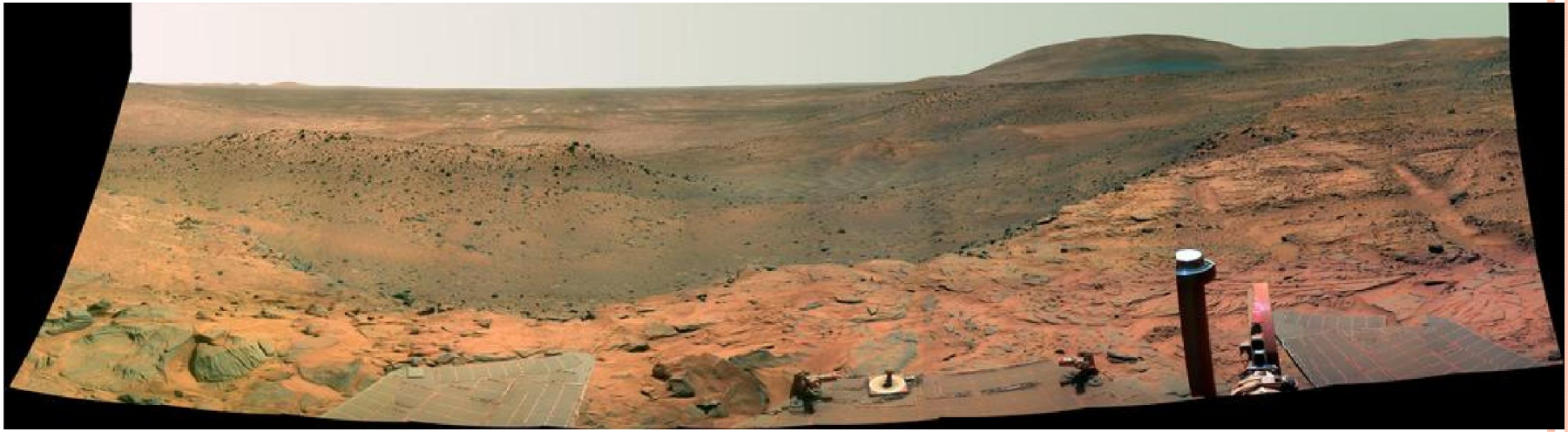
http://www.nytimes.com/2010/10/10/science/10google.html?ref=artificialintelligence

Interactive Games: Kinect

- Object Recognition: http://www.youtube.com/watch?feature=iv&v=fQ59dXOo63o
- Mario: http://www.youtube.com/watch?v=8CTJL5lUjHg
- 3D: http://www.youtube.com/watch?v=7QrnwoO1-8A
- Robot: http://www.youtube.com/watch?v=w8BmgtMKFbY
- 3D tracking, reconstruction, and interaction: http://research.microsoft.com/en-



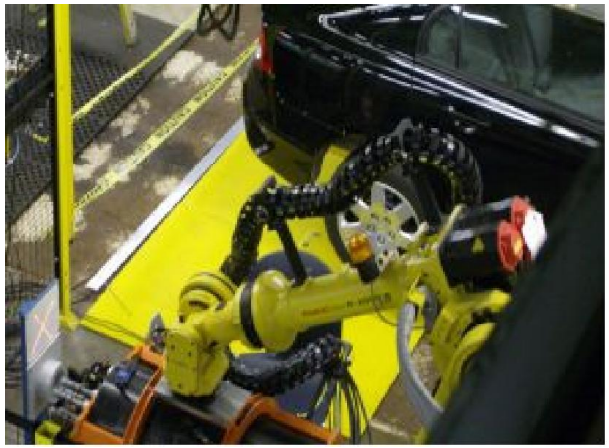
VISION IN SPACE



NASA'S Mars Exploration Rover Spirit captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

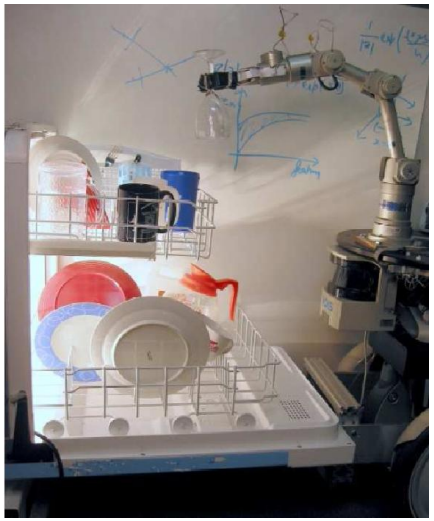
From wiki

INDUSTRIAL ROBOTS

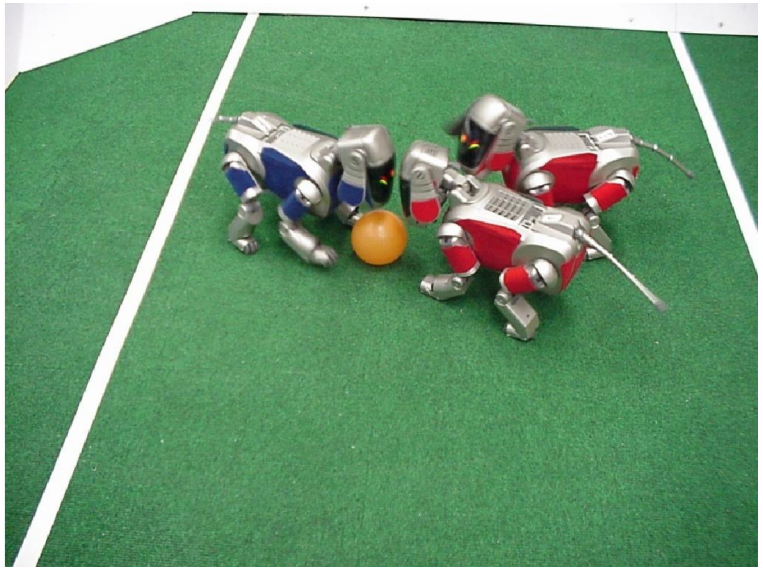


Vision-guided robots position nut runners on wheels

Mobile robots

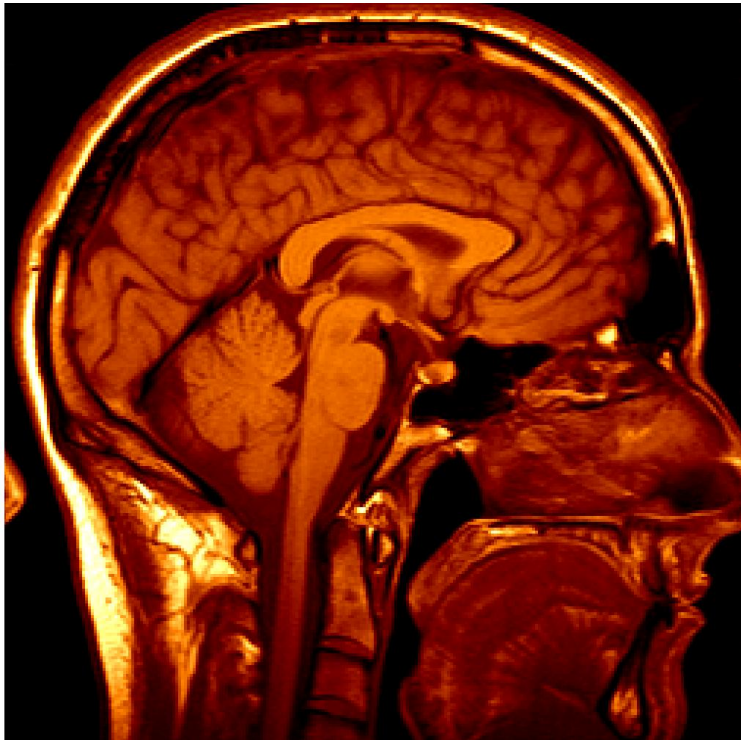


Saxena et al. 2008
STAIR at Stanford



http://www.robocup.org/

Medical imaging



3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT

INTRODUCTION

- Object Detection is having an important role in maximum of the machine vision applications, being it vehicle navigation, surveillance, autonomous robot routing, industrial robots, human machine interaction etc.

- Object position and orientation could be the other requirement in object detection.

Object Detection

- Object detection can be defined as the process of segmenting an object of interest from a video scene or image.

- *How to approach for object detection in a scene?*

Object Detection

- Characteristics of objects used to classify object of interest from rest other objects in the scene.

- Object Features
  - Object shape/size/texture features
    - For human detection, these could be face/pedestrians features
  - Object motion features

Object Detection

- Various Approaches
  - Viola-Jones Approach
  - Frame-Differencing Approach
  - Skin Color Modeling
  - Optical Flow Technique

VIOLA-JONES ALGORITHM

- Viola-Jones algorithm [1] is used for human detection.
- This algorithm uses Haar-like features and a Cascaded classifier.
- Haar-like features
  - Feature is the encoded information characterizing a class of an object.
  - Haar-like features utilizes texture information of object therefore features are only created for face part of human body for human detection.

VIOLA-JONES ALGORITHM

- Haar-like features

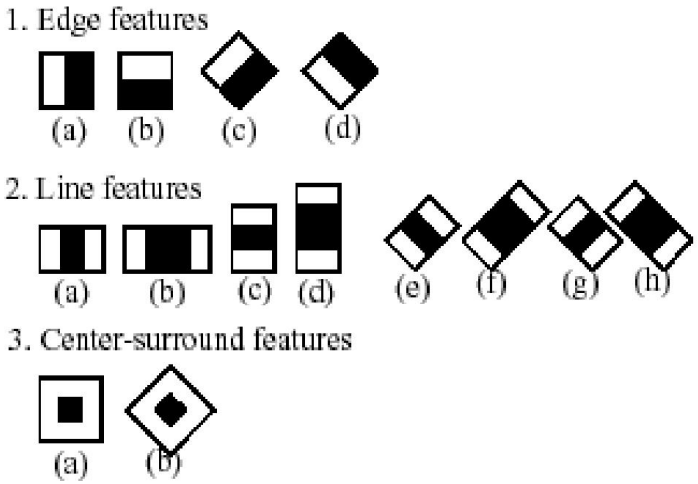


Figure: Haar-like features [2]

- Computation of these features does not add extra complexity as explained in [1, 3].
- Features generated are very large in number, eg. over 117,000 rectangle features for a 24x24 sub window. [3]

VIOLA-JONES ALGORITHM

- A very large quantity of feature makes processing of feature-based image more complex compared to pixel-based image.
- But, a Cascaded classifier solves this problem.

- Cascaded classifier
  - Cascade classifier is a cascade of a number of weak classifiers.
  - A weak classifier is designed to select the single rectangle feature that best separates negative & positive samples.

$$h_j(x) = \begin{cases} 1, & p_j f_j(x) < p_j \theta_j \\ 0, & otherwise \end{cases}$$

VIOLA-JONES ALGORITHM

- Cascaded classifier

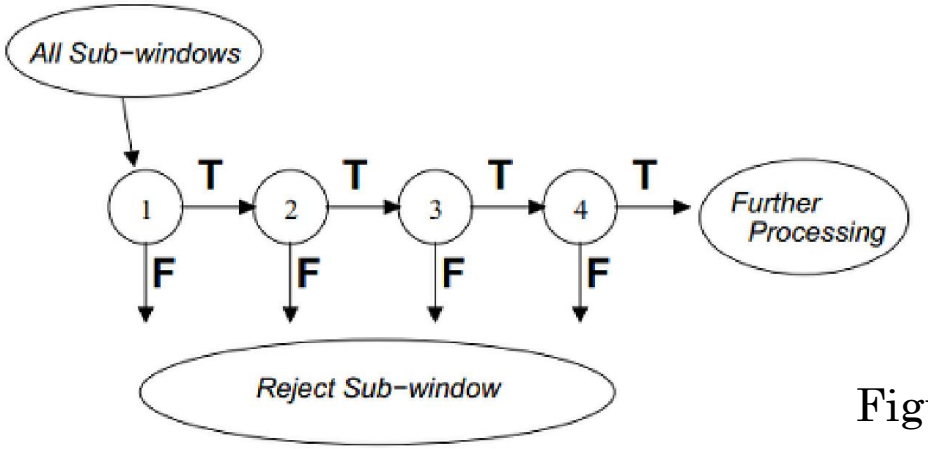


Figure: Cascade of Classifier [3]

- A weak classifier at every stage rejects a number of sub-windows, boosting the next stages of classifier [3].
- This speeds up the detection rate.

VIOLA-JONES ALGORITHM

- Experimental Results

Figure: Detected Face from sampled frame

- A geometric adjustment to the aspect ratio of height and width is done to mark the whole body.

VIOLA-JONES ALGORITHM

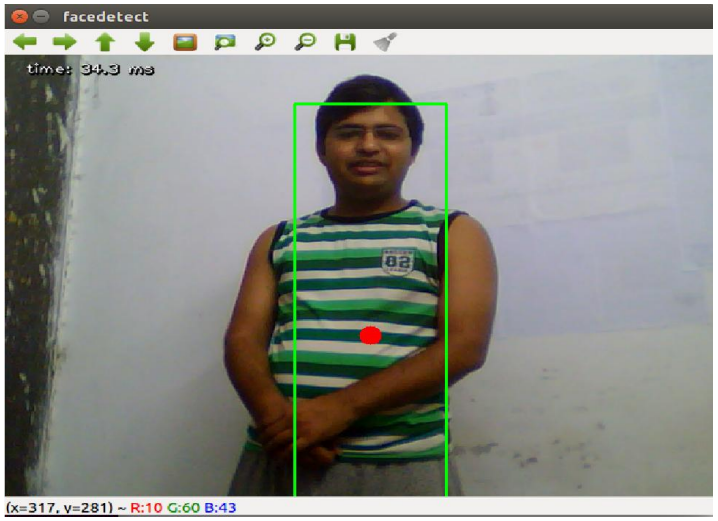


Figure: Left movement

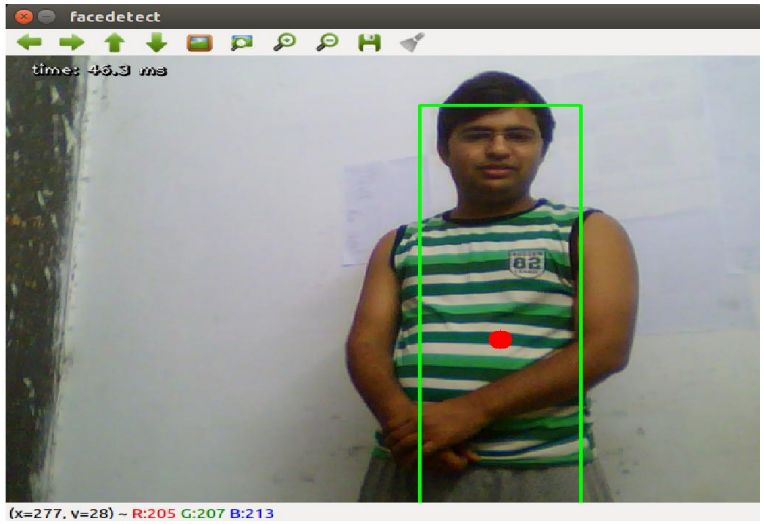


Figure: Right movement

- The red dot denotes the centroid of object.

Frame-Differencing Approach

- This approach exploits the motion features of the object.
- It is applied on continuous sequence of frames of any video captured by a static camera.
- The difference between pixels which are not in motion give zero as result and if the value is non zero then object corresponding to those pixels is in motion.

FRAME-DIFFERENCING APPROACH

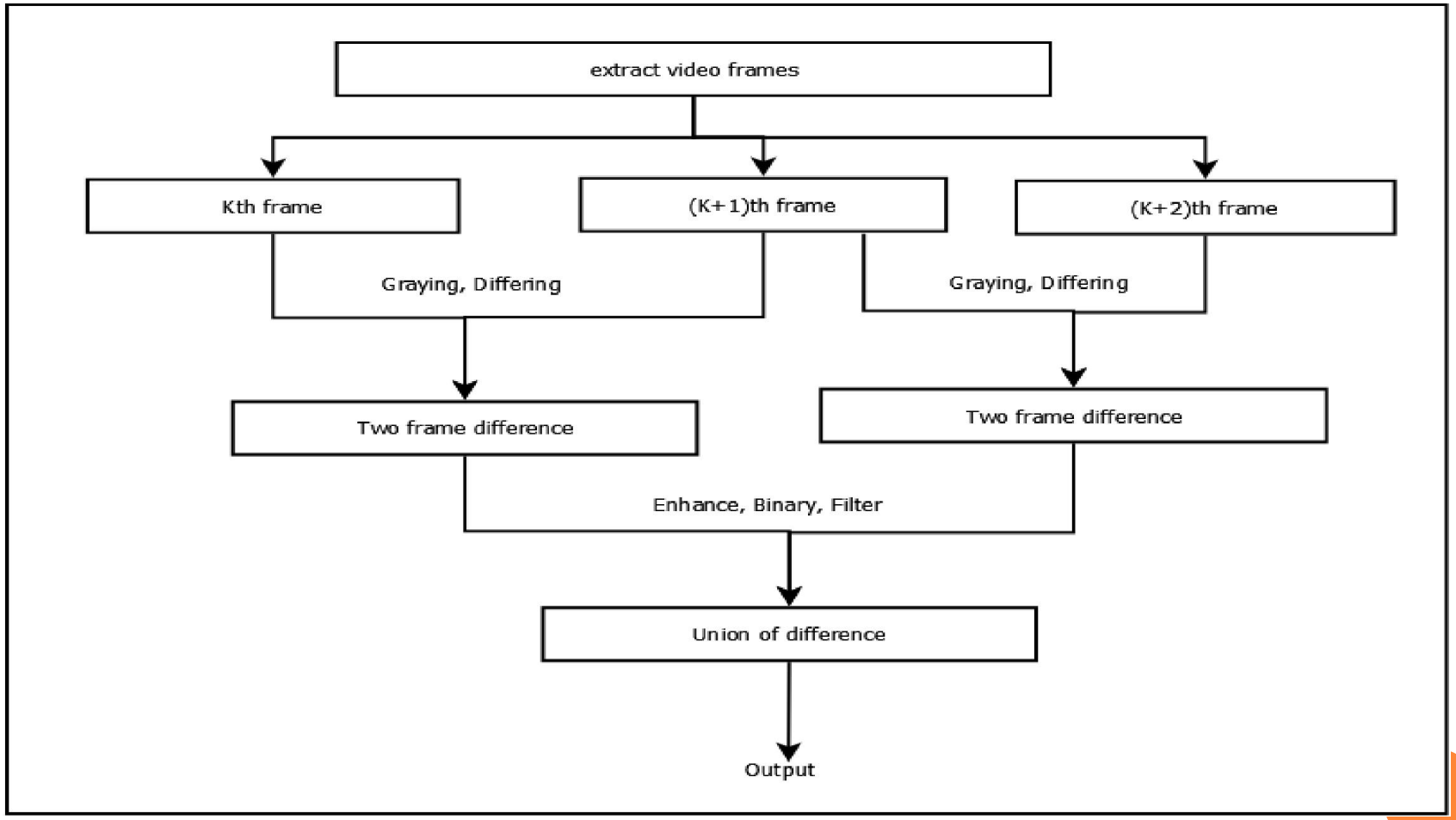


Figure: Flow chart of three frame difference

Frame-Differencing Approach

a) First frame

b) Second frame

c) Moving car detected

Figure: Result of frame differencing to detect moving car

Problems With Two Approaches

- Background subtraction & Frame-differencing techniques [4, 5, 6, 7, 8]
  - Constraints:
    - more prone to background noise and movements in background;
    - affects the correct detection of object [7]

Figure: Results of Frame-differencing approach

Problems With Two Approaches

- Viola-Jones and other classifiers based techniques
  - Use features as Haar / Hog etc. with some classifiers. [9, 10, 11, 12, 13]
  - Texture of human face is the basis for these features.
  - Constraints:
    - solely applicable for face portion
    - algorithm doesn't work, if face gets turned or tilted [11, 12]

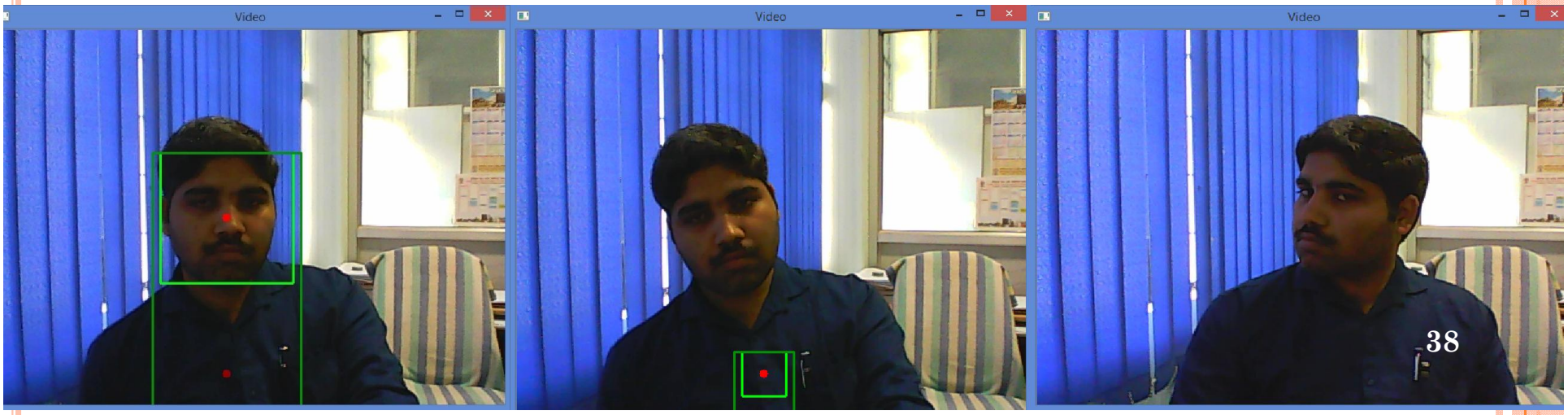


Figure: Results of Viola-Jones approach

Skin Color Modeling

Skin Color Modeling based Detection

- Skin Color Modeling [14-34]
  - Procedure to model a classifier that can classify skin and non-skin pixels in an image.
  - Skin color, mostly visible on face, is a feature that fairly distinguishes human beings from other classes of objects.
  - Skin color occupies a specific range of intensities in each plane of a color space.
  - Image pixels lying in this range are recognized as skin pixel.

Skin Color Modeling based Detection

- Classification:
  - When a classifier "$C_{x,y}$" is applied on image "$I_{m,n}$", the classified output is a binary image.

$$I_{m,n}(C_{x,y}) = \begin{cases} 1 & if\ I_{mn} \in skin\ region \\ 0 & otherwise \end{cases}$$

  - x, y are the random variables used to denote planes of color space.

- Recognized skin pixels form a blob, which is one candidate to the human face.

Skin Color Modeling based Detection

- Geometrical approximation of the blob w. r. t. human face is done for confirming a blob as human face.
- Aspect ratio of face width to face height is used for geometrical approximation.
- Matched blob is used to mark the position of human in the image.

Skin Color Modeling based Detection

- Challenges:
  - Range of skin color varies based on age, ethnicity and geographical location.
    - **An algorithm** which is able to cover skin color range variations.
  - Effect of ambient illumination on image disturbs skin color range.
    - Effects of illumination can be reduced by choosing **appropriate color space**.

Skin Color Modeling based Detection

- Color space:
  - YCbCr color space

$$Y = 0.299R + 0.587G + 0.114B$$

Luminance Component

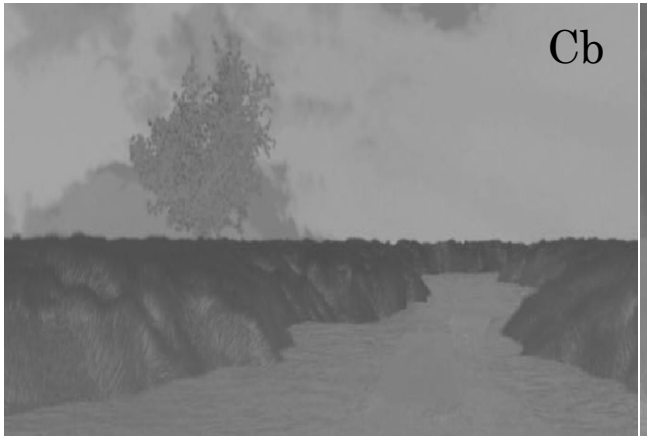
$$Cr = R - Y$$
$$Cb = B - Y$$

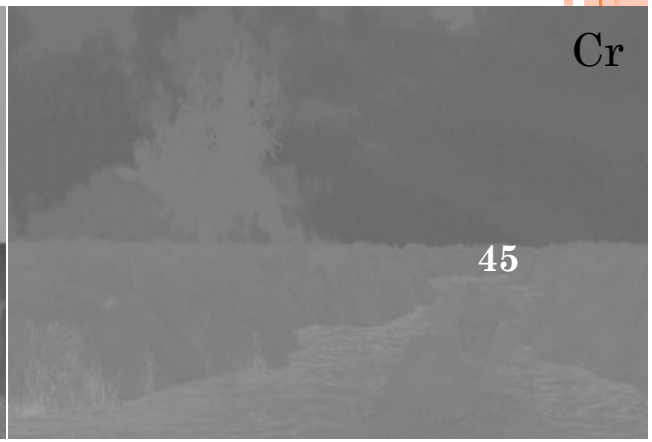
Chrominance Component

  - Cb, Cr components are resistant to illumination effects. [14, 16]

Skin Color Modeling based Detection

- RGB image and corresponding Y, Cb & Cr transformed planes

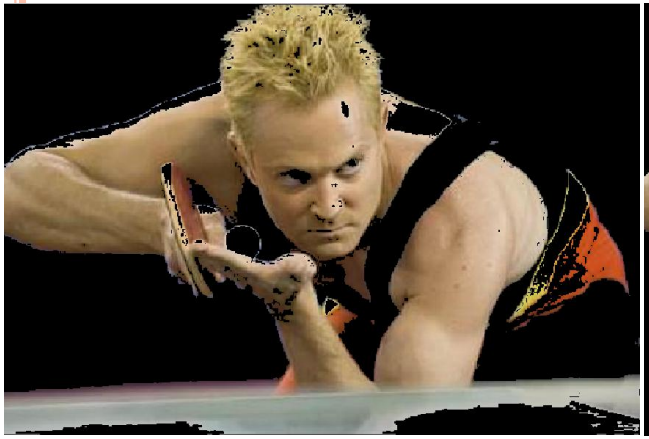


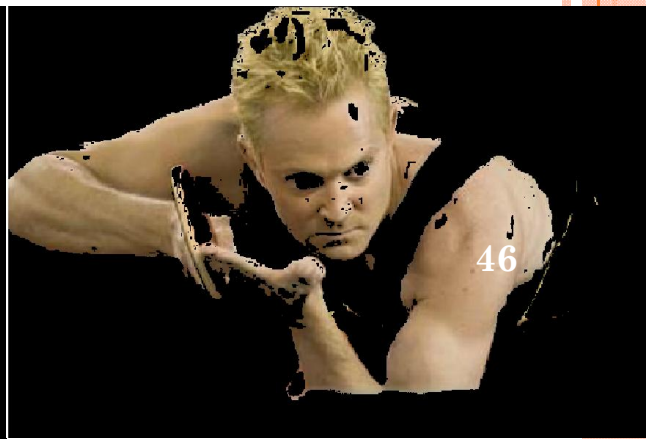


Skin Color Modeling based Detection

- A test using matlab thresholder app

R[70 - 240], G[40 - 210], B[20 - 190]
H[20 – 55 deg.], S[0.15 – 0.7], V[0.3 – 0.9]
Y[65 - 200], Cb[95 - 110], Cr[135 – 150]





Skin Color Modeling based Detection

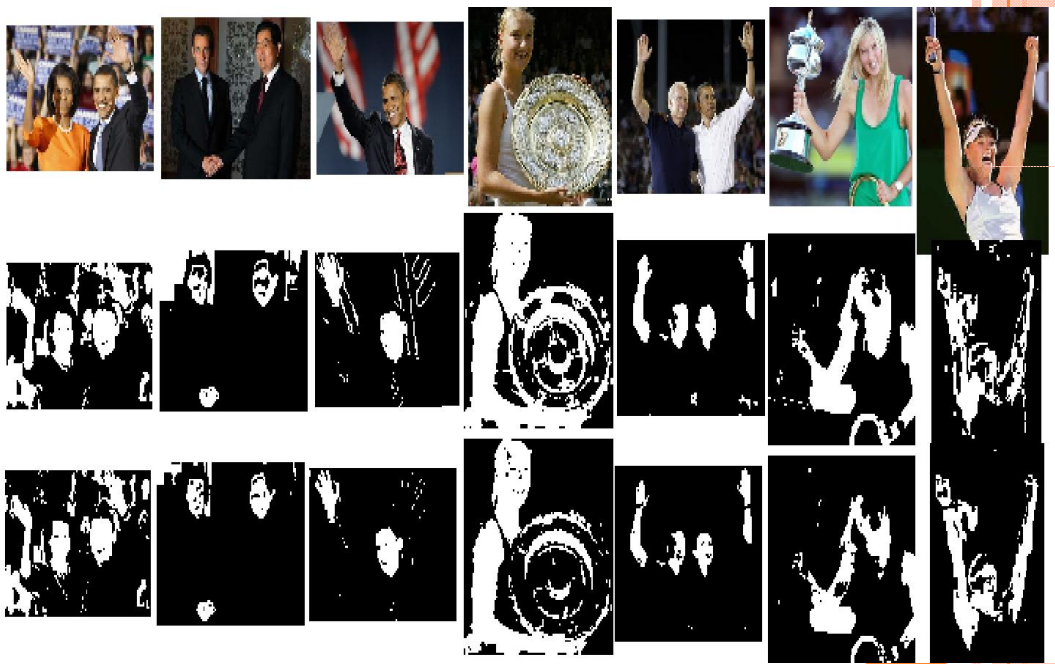
- Two Proposed Method

Gaussian Elliptical Fitting based Skin Color Modeling for Human Detection

ILUT based Skin Color Modeling for Human Detection

Skin Color Modeling

- Experimental Results



Skin Color Modeling

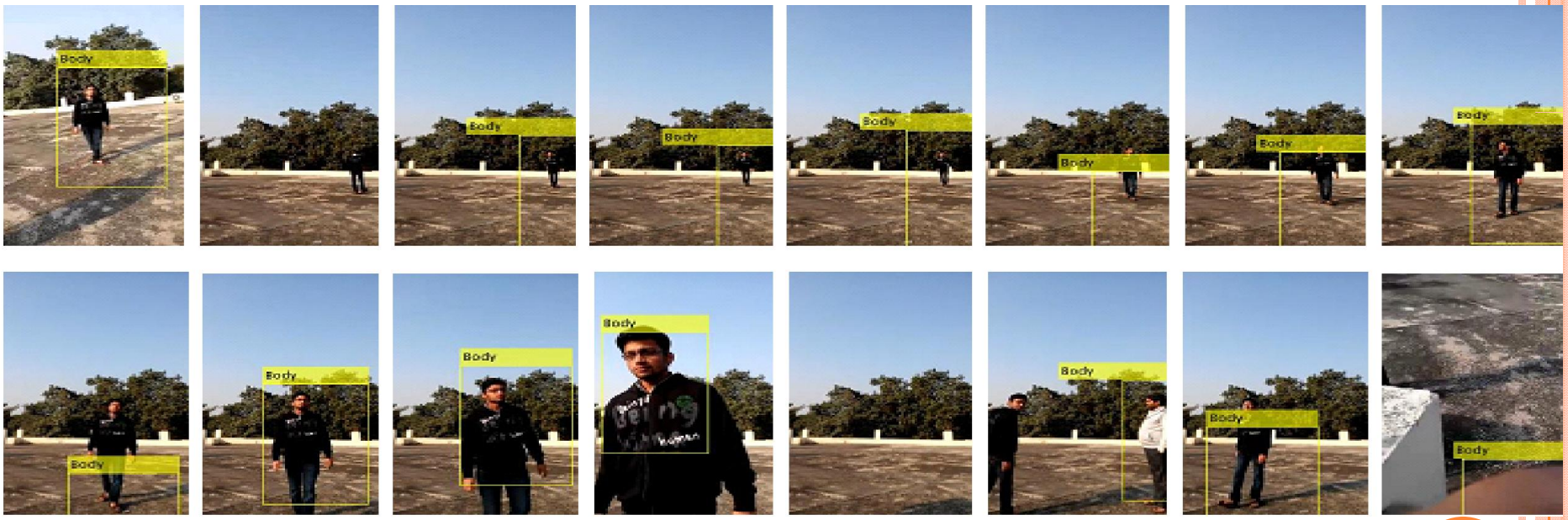


Figure: Image results of human detection by the proposed approach

Skin Color Modeling

- Conclusion
  - Skin color modeling based human detection overcome the problems seen with Frame-differencing approach and the Viola-Jones approach.
  - Skin Color Modeling based approaches face problem if human face is covered/occluded.
  - Human detection is not possible since skin features are not available.
  - This problem is prevalent for scenes where surveillance space (beyond border fences) is not in our control.
  - To overcome this problem, motion features are then used in next proposed approach for human detection.

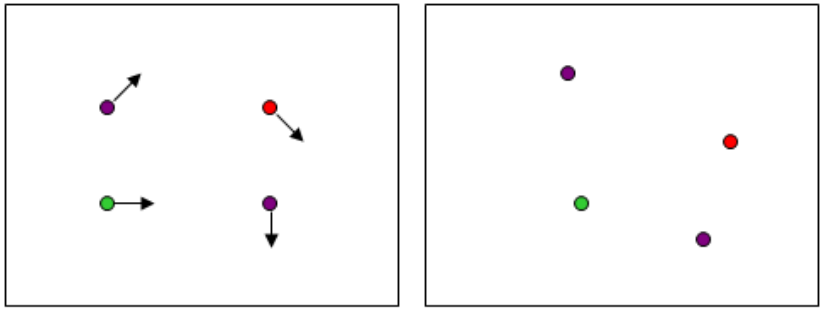
Optical Flow based Detection

- Optical Flow technique [41- 44]

- Proposed Approach

Optical Flow based Detection

- Optical Flow technique
  - It calculates the velocity of all the points in the image.
  - $[u, v]$ is the velocity vector in two dimensions of coordinate plane.
  - Figure given below shows the movement of points with their direction of movement.



H(x, y), the previous frame

H'(x, y), the next frame

Figure: Optical Flow of points in two consecutive frames

Optical Flow based Detection

- Optical Flow technique
  - The key assumptions taken for calculation of velocity vector [u, v] are:
    - *Color / Brightness constancy*: The intensity of a point p remains same during the movement in two frames.

      $$f(x(t), y(t), t) = c$$

    - *Temporal Persistence*: This means slow motion of object i.e. point p doesn't move very far.

      $$I_x u + I_y v + I_t = 0$$

    - *One equation and two unknowns; solution to this is proposed by Lukas-Kanade [4, 5] by making another assumption of 'Spatial Coherence'.*

Optical Flow based Detection

- Optical Flow technique
  - The third key assumption is:
    - *Spatial Coherence*: Pixel's neighbors have same [*u, v*]
    - A set of simultaneous linear are derived for a group of neighboring pixels.
    - Least Square method [45, 46] is used for solving equations for '*u*' & '*v*'.

Optical Flow based Detection

- Proposed Approach
  - **Algorithm: Human Detection**
  - **Input:** Sampled Frames
  - **Output:** Detected Human Being (marked with yellow rectangle)
    - *Optical Flow* (1st Frame, Next Frame) calculation of velocity vectors of moving points
    - Noise filtering and human detection on the binary output of Optical Flow. This is done by checking connected components and performing thresholding on connected component size.

Optical Flow based Detection

- Proposed Approach
  - .....
  - .....
  - Connected Components are the group of pixels where each pixel of group is well reachable to other.
  - Thresholding is done on size of connected components based on number of pixels. {An experiment for deciding threshold limit is done}.
  - The human being detected is marked with a yellow rectangle in the final output.

Optical Flow based Detection

- Proposed Approach
  - **Algorithm: Optical Flow**
  - **Input:** Two Images (Image1, Image2)
  - **Output:** Binary image showing velocity vectors of moving points
  - **Terminologies:** *Ix* & *Iy* are the horizontal and vertical components of displacement of a moving point respectively, *It* is the time domain difference of two images.
    - Convolute Image1 with matrix M1 to find *Ix*.

$$M1 = \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}$$

Optical Flow based Detection

- Proposed Approach
  - .....
  - .....
  - Convolute Image1 with matrix M2 to find *Iy*.

    $$M2 = \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix}$$
  - Convolute Image1 with matrix M3 and Image2 with matrix M4, to find *It*.

    $$M3 = \begin{bmatrix} 1 & 1 \end{bmatrix} \qquad M4 = \begin{bmatrix} -1 & -1 \end{bmatrix}$$
  - Lukas-Kanade algorithm is applied by considering a block of 3x3 pixels to have a system of simultaneous linear equations in 2 variables. *u* & *v* components of velocity vector of a moving point are the solution of Lukas-Kanade algorithm.

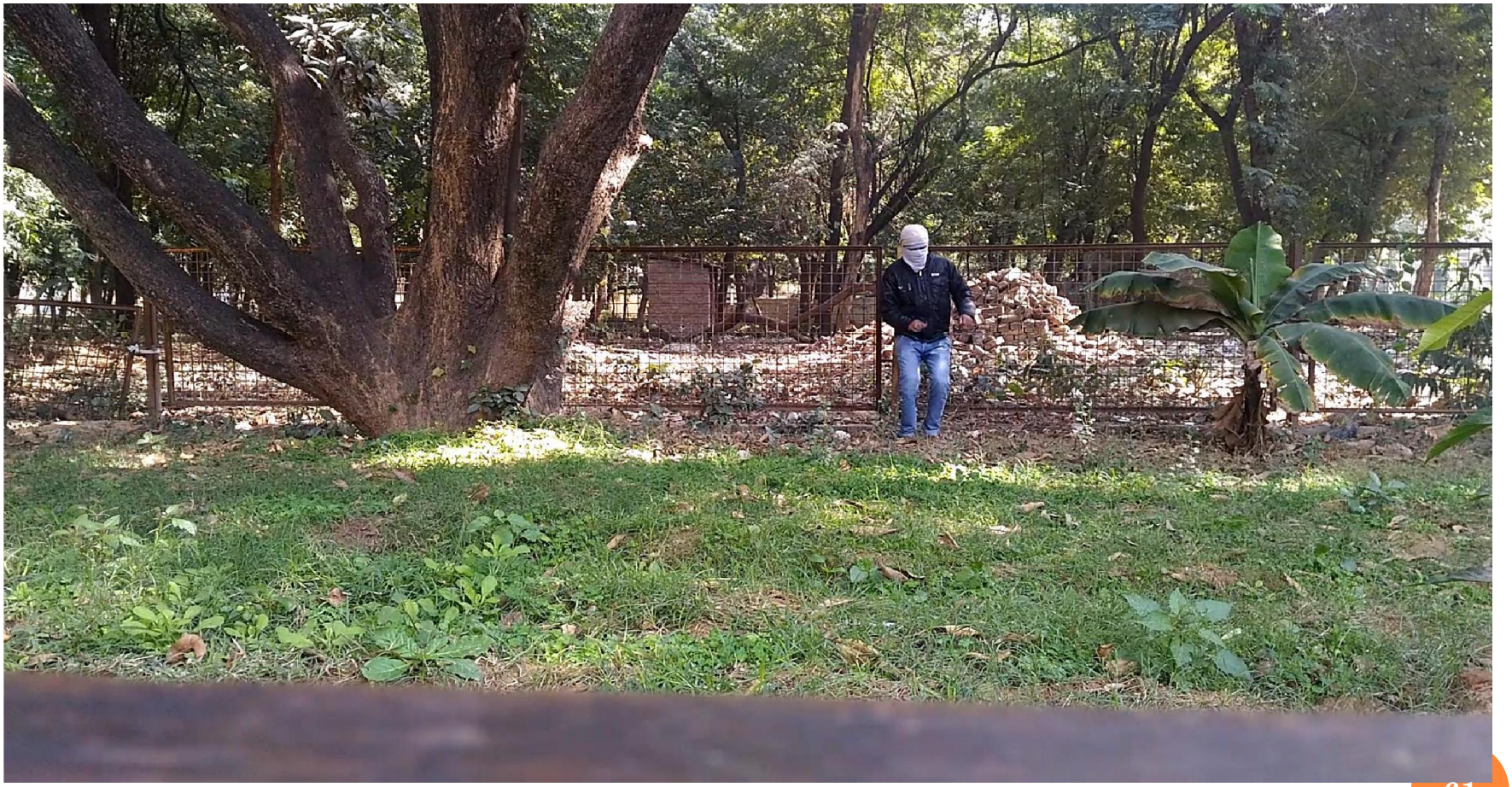
Optical Flow based Detection

- Proposed Approach
  - .....
  - .....
  - Previous step is repeated for all points in the Image1 having non-zero value in *It*.
  - Output is binary image showing velocity vectors at only moving points, with white. Rest all static points are black.

Optical Flow based Detection

- Experiment
  - Due to unavailability of actual border site surveillance video, a scene emulating almost the same behavior is created.

Optical Flow based Detection



- Metal grill imitates the actual border fence.

Optical Flow based Detection

- Experiment for Connected Components
  - First 20 sampled frames of the video are processed for different values of threshold.
  - Threshold or size of connected component is taken in terms of number pixels.
  - The number of frames out of twenty in which the human being is correctly detected for each value of threshold, is derived.
  - Table below presents the result

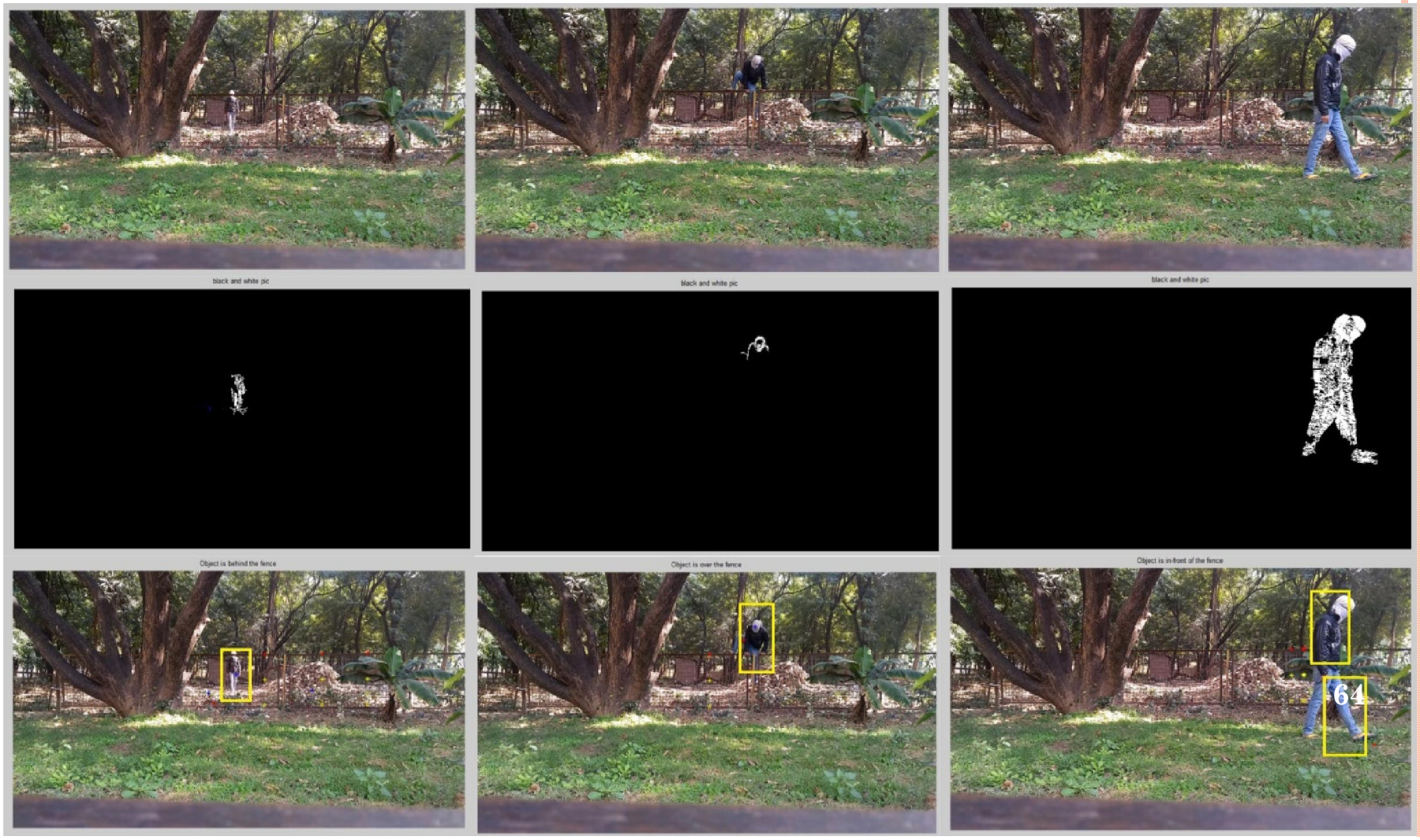
Threshold on Size of Connected Component	325	350	375	400	425	450	475	500	525
Number of frames with human correctly detected	4	9	14	**16**	15	12	8	2	0

Decision for Response/Combat

- Image Results (naked face)

Decision for Response/Combat

- Image Results (covered face)



Decision for Response/Combat

- Image Results

REFERENCES

[1] Viola, Paul, and Michael Jones. "Rapid object detection using a boosted cascade of simple features." *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*. Vol. 1. IEEE, 2001.

[2] Wilson, Phillip Ian, and John Fernandez. "Facial feature detection using Haar classifiers." *Journal of Computing Sciences in Colleges* 21.4 (2006): 127-133.

[3] Menezes, Paulo, José Carlos Barreto, and Jorge Dias. "Face tracking based on haar-like features and eigenfaces." *IFAC/EURON Symposium on Intelligent Autonomous Vehicles*. 2004.

[4] Power, P. Wayne, and Johann A. Schoonees. "Understanding background mixture models for foreground segmentation." *Proceedings image and vision computing New Zealand*. Vol. 2002. 2002.

[5] Heikkilä, Marko, and Matti Pietikäinen. "A texture-based method for modeling the background and detecting moving objects." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 28.4 (2006): 657-662.

[6] Ye, Jianhua, Tao Gao, and Jun Zhang. "Moving object detection with background subtraction and shadow removal." *Fuzzy Systems and Knowledge Discovery (FSKD), 2012 9th International Conference on*. IEEE, 2012.

[7] Weihua, Xiong, et al. "Moving object detection algorithm based on background subtraction and frame differencing." *Control Conference (CCC), 2011 30th Chinese*. IEEE, 2011.

[8] Qiang, Ganfeng, Lijun Yue, and Fengqi Yu. "A novel method of background subtraction for indoor surveillance." *Information Science and Technology (ICIST), 2014 4th IEEE International Conference on*. IEEE, 2014.

[9] Wu, Jianxin, et al. "C^4: A Real-Time Object Detection Framework." *Image Processing, IEEE Transactions on* 22.10 (2013): 4096-4107

REFERENCES

[10] Viola, Paul, and Michael Jones. "Robust real-time object detection."*International Journal of Computer Vision* 4 (2001): 51-52.

[11] Wu, Jianxin, et al. "Fast asymmetric learning for cascade face detection."*Pattern Analysis and Machine Intelligence, IEEE Transactions on* 30.3 (2008): 369-382.

[12] Wu, Jianxin, Christopher Geyer, and James M. Rehg. "Real-time human detection using contour cues." *Robotics and Automation (ICRA), 2011 IEEE International Conference on*. IEEE, 2011.

[13] Dalal, Navneet, and Bill Triggs. "Histograms of oriented gradients for human detection." *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*. Vol. 1. IEEE, 2005.

[14] Liu, Leyuan, et al. "Real-time skin color detection under rapidly changing illumination conditions." *Consumer Electronics, IEEE Transactions on* 57.3 (2011): 1295-1302.

[15] Tan, Wei Ren, et al. "A fusion approach for efficient human skin detection."*Industrial Informatics, IEEE Transactions on* 8.1 (2012): 138-147.

[16] Vezhnevets, Vladimir, Vassili Sazonov, and Alla Andreeva. "A survey on pixel-based skin color detection techniques." *Proc. Graphicon*. Vol. 3. 2003.

[17] Phung, Son Lam, Abdesselam Bouzerdoum, and Douglas Chai. "A novel skin color model in ycbcr color space and its application to human face detection."*Image Processing. 2002. Proceedings. 2002 International Conference on*. Vol. 1. IEEE, 2002.

[18] Chen, Wei, Wang, Ke and Li, Ming. "Skin color modeling for face detection and segmentation: a review and a new approach." *Multimedia Tools and Applications* 74.1 (2014): 321-345.

REFERENCES

[19] Peer, Peter, and Franc Solina. "An automatic human face detection method." (1999): 122-130.

[20] Batagelj, Borut, Franc Solina, and Peter Peer. "15 seconds of fame: an interactive, computer-vision based art installation." *Proceedings of the 12th annual ACM international conference on Multimedia*. ACM, 2004.

[21] Chen, Wei-Che, and Ming-Shi Wang. "Region-based and content adaptive skin detection in color images." *International journal of pattern recognition and artificial intelligence* 21.05 (2007): 831-853.

[22] Menser, Bernd, and Mathias Wien. "Segmentation and tracking of facial regions in color image sequences." *Visual Communications and Image Processing 2000*. International Society for Optics and Photonics, 2000.

[23] Vadakkepat, Prahlad, et al. "Multimodal approach to human-face detection and tracking." *Industrial Electronics, IEEE Transactions on* 55.3 (2008): 1385-1393.

[24] Rahman, Mohammad T., and Nasser Kehtarnavaz. "Real-time face-priority auto focus for digital and cell-phone cameras." *Consumer Electronics, IEEE Transactions on* 54.4 (2008): 1506-1513.

[25] Gonzalez, Rafael C., Richard Eugene Woods, and Steven L. Eddins. *Digital image processing using MATLAB*. Pearson Education India, 2004.

[26] Zarit, Benjamin D., Boaz J. Super, and Francis KH Quek. "Comparison of five color models in skin pixel classification." *Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, 1999. Proceedings. International Workshop on*. IEEE, 1999.

[27] Jones, Michael J., and James M. Rehg. "Statistical color models with application to skin detection." *International Journal of Computer Vision* 46.1 (2002): 81-96.

Thank You

One Week Workshop

On

Image and Video Analytics (WIVA - 2018)

Date: 29th March, 2018 – 02nd April, 2018

Objective

This course covers introductory and advanced knowledge of Image and Video Processing. Various problems of the Computer Vision domain will be explored. Vision enables computer system to see and interpret the image and videos for the objects present in it. This makes computer and machines to understand the surroundings and behave accordingly. We can examine its role in many applications as surveillance, driver-assistance, driverless cars, human-machine interaction, intelligent agents, medical imaging and many more. Any kind of objectives achieved manually using cameras can be approached as an automatic solution with image and video processing fundamentals. This is really helpful to the mankind as the world is looking forward for automation in every domain of life. Since computer vision is a multidiscipline concept so workshop will also be covering some machine learning aspects of vision in addition to the image and video processing concepts. We also aim at imparting knowledge of current research spaces and challenges in image and video processing domain. MATLAB and Python with OpenCV will be used as the tools for practical handshaking.

Course Contents

- **Image and Video Acquisition and Representation techniques**
- **Image and Video Signal Processing**
- **Feature Extraction and pre-processing.**
- **Image Enhancement and Filtering techniques**
- **Image Segmentation**
- **Image/Video Annotation, Search and Retrieval**
- **Video Segmentation, Object Detection & Tracking**
- **Visual perception, human visual system**
- **Computer and Machine Vision**
- **Vision for Surveillance**
- **Medical Imaging**
- **Human Machine Interaction**

Resource Persons

- **Professors from IITs, IIITs, NITs**
- **Experts from Industries**
- **Faculty from MNNIT Allahabad**

Important Date

Registration Deadline : 15th March 2018

Registration Fee

U.G. Students 1000/-
P.G. Students 1500/-
PhD Students 2000/-
Faculties and Others 2500/-

Contact

Coordinator: +91-9359133388
E-mail: 2018wiva@gmail.com
Student Coordinator: +91-7004394850, +91-9889092231

Course Coordinators

Dr. Dushyant Kumar Singh
Dept. of CSE, MNNIT Allahabad
Email: dushyant@mnnit.ac.in

Dr. Rajitha B.
Dept. of CSE, MNNIT Allahabad
Email: rajitha@mnnit.ac.in

Department of Computer Science & Engineering
Motilal Nehru National Institute of Technology Allahabad
Allahabad-211004, Uttar Pradesh, India