

AVI SINGH

Fourth Year Undergraduate, IIT-Kanpur

✉ avisingh599@gmail.com 📞 **+(91) 8853544535** 🌐 <https://github.com/avisingh599> 🌐 <http://avisingh599.github.io>

RESEARCH INTERESTS

Computer Vision, Machine Learning, Robotics

EDUCATION

Indian Institute of Technology Kanpur | GPA: **9.4/10** (after six semesters) • Bachelor of Technology (expected July 2016)
Major: Electrical Engineering • Minor: Computer Science (Artificial Intelligence)

SCHOLASTIC ACHIEVEMENTS AND AWARDS

- 2014 Awarded an **A* grade**, for exceptional performance in **Undergraduate Project - 1**
- 2012-2013 **Academic Excellence Award**, IIT Kanpur (Awarded to 60 students out of 840).
- MAY 2012 Secured All India rank **387** in IIT-JEE 2012 out of **0.5m** candidates
- MAY 2012 Secured All India rank **345** in AIEEE 2012 out of **1.2m** candidates
- JAN 2012 **Top 1%** in National Physics Olympiad

PUBLICATIONS

[Recurrent Neural Networks for Driver Activity Anticipation via Sensory-Fusion Architecture](#) <http://arxiv.org/abs/1509.05016>
Ashesh Jain, **Avi Singh**, Hema Koppula, Shane Soh, Ashutosh Saxena
IEEE International Conference on Robotics and Automation (ICRA), Stockholm, Sweden, 2016 (**under review**)

RESEARCH INTERNSHIP AT CORNELL UNIVERSITY

[Brain4Cars - Anticipating Maneuvers via Learning Temporal Driving Models](#) MAY 2015 - JULY 2015
under Prof. Ashutosh Saxena, Robot Learning Lab, Department of Computer Science, Cornell University www.brain4cars.com
Brain4Cars takes input from driver-facing and road-facing cameras, and predicts what maneuvers the driver is going to perform in the next few seconds. My contribution to the project was in the following areas:

- Revamped the vision pipeline of the project by incorporating a Constrained Local Neural Field model for enhanced head pose tracking of the driver.
- Implemented Gaussian Mixture Model-based initialization, and LBFGS optimization for training Autoregressive Input Output Hidden Markov Models.
- The above changes improved the performance of the AIOHMM-based anticipation system from a Precision/Recall of 77.4/71.2 to 86.7/78.2.
- Further testing on a Recurrent Neural Network with Long Short Term Memory (LSTM) cells increased the performance to a Precision/Recall of 90.5/87.4.

RESEARCH EXPERIENCE AT COMPUTER VISION LAB, EE, IIT-KANPUR

[Visual Odometry for Ground Vehicles](#) AUG 2014 - DEC 2014
under Prof. KS Venkatesh, Computer Vision Lab, Department of Electrical Engineering, IIT-Kanpur
Two **stereo** approaches were implemented and evaluated on KITTI odometry benchmark:

- Jiang2014: Model Based ICP: 3D points triangulated from stereo data, inliers detected via the use Iterative Closest Point Algorithm that uses a 1-DOF motion model for initial estimate. Efficient PnP Algorithm is then used on the selected inliers to obtain the final rotation and translation
- Howard2008: Inlier detection using an assumption of scene rigidity. Problem reduced to finding the maximum clique in a graph, solved using a heuristic. Levenberg-Marquardt used for minimizing the reprojection error on the selected inliers. Implementation available on [Github](#).

A **monocular** visual odometry approach was also implemented:

- Utilizes Nister's five point algorithm for essential matrix computation. Implementation available on [Github](#).

[Estimation of Scene Flow from RGBD data](#) MAY 2014 - JULY 2014
under Prof. KS Venkatesh, Computer Vision Lab, Department of Electrical Engineering, IIT-Kanpur

- Designed and implemented two approaches for estimation of Scene Flow from RGB-D data, one using 'Total Least Squares Solution', the other utilizing 'Global Minimum Energy Solution'.
- Captured RGBD Data from a Microsoft Kinect using OpenNI and OpenCV libraries in C++, and qualitatively evaluated the results.

SELECTED PROJECTS

[DAAnt - Dental Automated Analysis Tool \(Computer Vision for intraoral cameras\)](#) JANUARY 2015
MIT REDX CAMP, under Dr. Hyunsung Park, Postdoc at Camera Culture, MIT Media Lab
We developed Computer Vision algorithms for diagnosis of dental diseased from images obtained from an intraoral camera. My contribution in the eight-member team was primarily in the following areas:

- **Stitching Images:** Used Affine SIFT to overcome the difficulties in feature matching due to changes in perspective, and then computed Homography with RANSAC to stitch the images.
- **Segmenting Every tooth:** Marker-controlled watershed transform was used to segment every tooth from the image. Both semi-automated and automated approaches were implemented and compared.

APRIL 2015

- Hidden Conditional Random Fields were used for human activity classification from RGBD sequences.
- Pose features were extracted from a skeleton detection pipeline
- A novel normalization operation was developed for reducing the effect of variance in body sizes of different subjects
- An accuracy of 71% was achieved on a reduced version of MSR Daily Activity 3D Dataset (6-class classification)

DECEMBER 2013

- Geometric Triangulation was used to determine the pose (modelled as a set of random variables with Gaussian Distribution) of a robot, from information obtained from noisy landmarks.
- A Microsoft **Kinect** was used to identify the landmarks (using color histogram based models), **CAMshift** algorithm was used to track these landmarks, and bearing measurements were calculated using the depth data.
- An error model for the data obtained from a Kinect was used along with the Error Propagation Law to arrive at the uncertainty in the final pose computed using the Geometric Localisation Algorithm.

MARCH 2014

[Github]

- Implemented a 32-point Discrete Hilbert Transform in **Verilog**, making use of the **Fast Fourier Transform** in the intermediate stages.
- Won **2nd position** in FPGA Design, Techkriti 2014.

Hardware Hacks
@ Electronics Club

Mathematics	Electrical Engineering and Computer Science	Other
Probability and Statistics	Machine Learning for Computer Vision	Engineering Graphics
Linear Algebra	Digital Signal Processing	Applied Game Theory
Multivariate Calculus	Data Structures and Algorithms	Intro to Psychology
Ordinary/Partial Differential Equations	Communication/Information Theory	Art of Video Making

Programming Languages	C, C++, MATLAB, Python	Microcontrollers	Atmel ATmega, dsPIC, Arduino
Software Packages	OpenCV, ROS, Point Cloud Library (PCL)	Imaging Systems	Microsoft Kinect, PST-110
Operating Systems	GNU/Linux (Ubuntu), Microsoft Windows	Other Hardware	Raspberry Pi, Beagleboard, Servos, IMUs
Other	Git, Bazaar		

Electronics Club Coordinator, Science and Technology Council	2014-2015
--	-----------

2014-2015

- Floated, mentored and ensured the completion of **nine summer projects** including a **3D Laser Scanner**, **A Video Surveillance Robot**, Conway's Game of Life simulation using FPGAs, Fast Fourier Transform on FPGA, An accelerometer based fitness and sleep tracker with accompanying Android App, a Surveillance system with **face recognition**, and a Laser Tag system.
- Leading a team of **16 secretaries** and handling a budget of **Rs.76,000** to organize lectures, workshops, competitions, and another **Rs.4,74,000** for funding projects and for participation in external events.
- Lectures attended by 400+ people, workshops attended by 200+ people, and participation of 100+ people in Takneek (intra-IIT Kanpur technical festival) Electronics competitions.

2015 - Present

I have given a talk about my work, and attended lectures given by others in the various sub-fields of Machine Learning.

2013 - 2014

Designed and verified the problem statement for the event Electromania. Prepared sample codes and tutorials for the participants.

2013 - 2014

Assisted in organisation of lectures, workshops, tutorials, and maintenance of club.

2013 - Present

Assisted in the organisation of various counselling service activities such as the Orientation Program and helped a group of six freshers in settling in the new college environment.