Cover Letter

Vikas Thamizharasan **vikas.tmz@gmail.com**

The paramount step in creating Artificial Intelligence is to teach computers to understand our world as intuitively as humans do. This dream is rapidly transforming into reality with the advancements in Computer Vision (CV), Machine Learning (ML), Natural Language Processing (NLP) and Computer Graphics, and in some cases outperforming human capabilities. It is my dream to be an integral part of this change. My research interests lie in the intersection of Image Processing, Computer Vision, Computer Graphics and Deep Learning. I wish to tackle problems on 3D shape and pose estimation, Depth Estimation, Image-based modeling and rendering, Generative models, Unsupervised learning (specifically disentangled representation learning), Geometric DeepLearning, Neural Architecture Search (NAS) and graph classification. I am seeking for an internship before I pursue my Masters in 2020.

Education

I obtained my Bachelor's in Computer Science and Engineering from the International Institute of Information Technology (IIIT), Hyderabad. My decision to join IIIT-H was due to it being a premier research and Tier 1 institute, the program curriculum and its vicinity within the thriving tech community of Hyderabad. The courses I took in my freshman and sophomore years *[4,CSE']*, introduced me to the fundamental concepts of Computer Science and provided me with a platform to learn and implement the underlying theories in various facets of Computer Science.

I had little idea about what research in this field entailed until in my final semester of my sophomore year, I took the Digital Signal Analysis and Applications course under *Prof. Vineet Gandhi* and the Computer Graphics course under *Prof. P. J. Narayanan*. I was enthralled by concepts like fourier analysis, signal compression, algorithms in triangulation, geometric query and ray tracing and the variety of applications built on them. To expand my knowledge base in this field, I enrolled for courses such as Digital Image Processing (DIP), Statistical Methods in Artificial Intelligence (SMAI), Computer Vision (CV), Linear Algebra and Distributed Systems. During these courses I got exposed to the intricacies of research through the evolution of ideas and theories.

As part of my course project in DIP under *Prof. Vineet Gandhi*, I proposed and worked on the topic of "***Virtual Garment Fitting from Single Image"***. The goal was given input images of the user and the target clothing item, re-target the respective clothing item onto the user using image segmentation, pose estimation and triangulation techniques. Though an exclusively application oriented project, through its development which included reading a great amount of literature, I was exposed to the difference between machine learning techniques including recent deep learning models and traditional ad hoc and hand-crafted methods based solely on heuristic knowledge. In this regard, I used methods ranging from graph cut *[5,Boykov et al']* to a parsing method based on nearest neighbor style retrieval by *[6, Yamaguchi et al']* for cloth parsing and similarly for feature point extraction, I used recent learning based techniques for pose estimation using *Part Affinity Fields (PAFs)* *[7, Cao et al']* to provide richer features. The re-targeting was performed via image morphing using delaunay triangulation. The project was selected as the winning idea for *Microsoft's code.fun.do hackathon* and I was given an opportunity to visit Microsoft Hyderabad campus for an all India showcase.

As part of my CV course under *Prof. Anoop Namboodiri*, I had to tackle the problem of object detection and classification as a group project. We opted for *YOLO* *[8, Redmond et al']* because it integrated detection and classification into one pipeline to achieve real time performance at high FPS and performed well in terms of precision and recall. The project also required us to create our own dataset as these classes are not available in the ImageNet dataset, but we utilized pre-trained weights to improve our classification.

In the penultimate semester of my undergrad, I worked on my bachelor thesis titled "***3D Model Reconstruction and Manipulation with a single image****" under Prof. Vineet Gandhi*. Over a period of two semesters we explored state of the art publications in conferences like ACM SIGGRAPH *[10, Chen et al']* and Eurographics *[11, Shtof et al']* that tried to tackle this problem and the improvements we could make upon them. It was a novel approach to generate 3D models from a single image using geometric primitives to infer geosemantic constraints followed by non-convex optimization fitting and model-to-image alignment. The result was an **interactive image editor** where objects could be manipulated in 3D space with the advantage of applying non-rigid transformations along with texture mapping and image illumination estimation to create realistic rendering, made available publicly *[12, code']*. Through this thesis I learned a lot about geometry-based methods in vision and mathematical optimization.

Having built an interest in Deep Learning, I wanted to explore other areas that have experienced significant impact with it development. Consequently during my final year I enrolled for the Information Retrieval and Extraction course under *Prof.*

Vasudeva Varma and worked on "***Deep learning approach for Query focused Abstractive Summarization***" as a team project. The goal was to generate summaries of text documents with respect to a certain query. We used Recurrent Neural Networks (RNNs) based encoder-decoder with a query and document attention model to learn temporal information and a diversity based attention model to alleviate the problem of repeating phrases in the summary. One application could be in search engines as featured snippets. During this project I realized the potential of RNN in taking advantage of temporal information.

Work Experience

Over the past few months, I was a research intern with the STARS Team at **INRIA** under the guidance of *Dr. Antitza Dantcheva* and *Dr. François Brémond*. In collaboration with Blue Manta (French startup), the internship was focused on performing depth estimation and generating low-dimensional face embedding for face analysis using deep learning techniques from raw data acquired using state of the art structured light and active infrared hardware. Unlike previous depth estimation techniques *[1, Eigen et al']* *[2, Fanello et al']* *[3, Zhang et al']*, we tackled this problem using GAN (Generative adversarial networks) trained on adversarial loss and a reconstruction loss for fine tuning. Due to the lack of datasets consisting of the above modalities, we created a synthetic and real dataset via a novel data generation pipeline. This allowed us to address the performance of deep network when trained on synthetic data, provide solutions to bridge the gap and also set benchmarks for Face Analysis tasks on different modalities. Our work has led to an ongoing publication. At INRIA, I got a chance to meet some of the leading European researchers in the field of ML, CV, Graphics and DL, attend an AI summit and volunteer for IEEE International Conference on Image Processing, Applications and Systems (IPAS) 2018.

Prior to joining INRIA, I was a Computer Graphics Engineer at **Carl Zeiss**, India where I worked on their Volume Rendering framework.

In the summer of 2017, I was selected for Google's fully funded open source program titled "***Google Summer of Code***" to work on the open source project ViSP developed by Lagadic (robotics lab) from INRIA, France mentored by Fabien Spindler and S. Trinh. ViSP (Visual Servoing Platform) is a modular cross platform library that allows prototyping and developing applications using visual tracking and visual servoing technique. The goal of this internship was to automate the creation of ViSP CAD model files from existing 3D formats and achieve perfect, lossless conversion *[9, Gsoc'17']*. This technique is vital in building tools to perform AR based tracking and interactive robotics. Through this internship, I learned valuable lessons on building tools that aid research work and subsequently real world applications.