INITIAL RESEARCH PROPOSAL FORM

(also referred to as 'Statement of Intent Form')

To be submitted by the researcher to the Institute Research Sub-Committee (IRC)

Research Title: **Using Scene Flow Analysis to improve the accuracy of SLAM**
Institute name MCAST: ICT Institute
Course / Programme: Top Up Degree in Software Development
Level and year of study Level 6 - 2018/2019
Main area of study being proposed: The main area of study is visual odometry in self driving cars. A tracking algorithm by the name of SLAM is commonly used in this area. This works by tracking visual features in a stream of images and maps them in a 3D space. This research will evaluate the accuracy of a new method that ignores tracks of non-stationary objects in images. Tests will be made on 'The KITTI Vision Benchmark Suite'. Various researchers have already submitted their algorithm's result which will allow us to compare our result with theirs.

Name of Researcher: Mr. Jason Formosa	**Researcher's I.D. Number:** 0221298M
Signature of Researcher	**Date of submission of Form** 04/04/2019
Name of Tutor (or Recommended Tutor): Mr. Roderick Vella	

Personal Motivation for the Choice of Research Theme.

A lot of research is being done on self-driving cars lately. Several companies are competing as to whoever can release a fully autonomous car first as they will have the lead in the market. Self-driving cars, however, depend on a particular sensor used to localize the vehicle in 3D space called 'LIDAR'. It's crucial that these sensors be as accurate as possible which is why these sensors cost a lot of money. This discourages independent researchers from contribute work and accelerating the development of self-driving cars. I aim to develop a system that uses off the shelf component and tries to retain the accuracy of commercial system as much as possible. This offers a cost effective solution for research to get into the field of self-driving cars.

Outline of Key Literature and Theoretical Framework or Propositions.

AI has become increasingly popular these last few years. A lot of research is being done on Neural Networks since computing power has increased dramatically in the past decade. A lot of research has been focused on CNNs (convolutional neural networks) since they can be greatly accelerated by current GPUs. CNNs are suited when you need to factor in the spatial information in the data to get the correct output. A good example of this are images. The spatial information in images is the way pixels are arranged next to each other.

In video analysis, optical flow analysis is a commonly used method to detect the pattern of motion of object in a scene relative to an observer, for example, a camera. Optical flow analysis is restricted to 2D motion. The motion of pixels is only tracked as moving horizontally or vertically in the camera frame. Scene flow analysis is used when you need to analyze the motion of objects in a 3D space from a sequence of images. This is harder to achieve than optical flow analysis since more data is required. One required attribute is the depth information in the camera frame, i.e. the distance between a specific pixel in the frame and the center of the camera sensor. One way to calculate this is by using two camera with a horizontal distance between them, similar to the way we humans use our eyes to get depth information. This works by calculating the disparity between the images of the two cameras at the same timestep, and by knowing the intrinsic parameters of the two cameras and the baseline between them, the depth can be extracted.

SLAM (Simultaneous Localization and Mapping) is a technique frequently used in robotics in order to localize an agent in a 3D space. There are different SLAM algorithms designed to work on different sources of input data, for example, 3D pointclouds obtained from LIDAR sensors, and a subset of SLAM called Visual SLAM which uses visual data as input, both mono and stereo. They work on the principle of tracking visual features and mapping their movement to estimate the motion of the camera.

In this research, a CNN will be used to provide the scene flow vector in the x, y, and z axis given two stereo image pairs at timesteps t and t+1. The movement of the vehicle recorded by an accelerometer will be subtracted from the movement of the pixels which should create a mask for all the stationary items in the scene. This mask will then be used to filter out all the erroneous tracking points obtained from the SLAM algorithm.

Significance of the Study.

The key objective of this study is to test whether removing erroneous feature points using scene flow data is a viable method to increase the accuracy of SLAM. Comparisons will also be made against SLAM using LIDAR pointclouds as input.

Hypotheses and/or Research Question/s

In this study, the assumption is being made that the accuracy of CNNs for scene flow analysis will be adequate for detecting erroneous feature points.

Is deep learning a viable way to estimate scene flow?

Is the addition of scene flow data enough to eliminate erroneous points?

How fast can this method be made in order to be a viable alternative to current production systems.

Target Participants and Research Methods for Data Collection and Analysis

For this study, The KITTI Vision Benchmark Suite will be used. It is a very popular dataset amongst researchers concerning self-driving cars. Researchers also upload the results for their algorithm on an online scoreboard on the dataset's website. This will provide other methods to evaluate against. Data used will be in the form of a sequence of images from a stereo camera mounted on the vehicle and the car acceleration from the accelerometer. As an evaluation metric, the translational and rotational deviation from the ground truth path will be used.

Ethical Considerations.

Refer to *guidance points below. You are also additionally required to read MCAST Document 074 'Research Ethics Policy and Procedure' that is available on the College website via link http://www.mcast.edu.mt/MainMenu/Full-TimeCourses/Rules,PoliciesandRegulations.aspx*

1. *Research shall be conducted in such a manner so as to avoid any psychological and physical harm to humans and animals and financial damage to organizations*
2. *Only the supervisor and examiners will have access to any data gathered.*
3. *Participants will remain free to withdraw from the study at any time without having to provide any reason. In the case of withdrawal, all the records and information collection will be deleted.*
4. *The participant, who is the sole proprietor of the data provided, is granting that such data would be processed for this study purposes only.*
5. *The data collection process will be a transparent process.*
6. *All transcriptions and/or electronic recordings reflecting the data collected, once exhausted, are to be deleted*
7. *Confidentiality, anonymity and data protection procedures are to be ethically abided by.*
8. *The researcher would provide a soft copy of the study to the participant, if required.*

Enter details here regarding possibility of issues regarding confidential personal data:

This study will not require the collection of any personal data.

Enter details here regarding possibility of physical harm:

All results from this study will be obtained through simulation on a computer. No physical harm should occur as a result of this study.

Enter details here regarding possibility of moral harm:

This research will not require any participants therefor direct moral harm is improbable. Some people are opposed to self-driving cars because they do not trust a machine with their lives, however this research is intended to increase the accuracy of current methods and therefor making them safer.

Enter details here regarding possibility of business harm:

This research shouldn't cause any harm to business.

Anticipated Contributions of the Study.

This study should serve as more of a learning experience in the field of image processing when it comes to self-driving cars. However this could raise interest amongst research to incorporate deep learning more in this field.

Dissertation Project Plan.

Early April: Submission of proposal.

Late April: Initial implementation of prototype.

Early-Mid May: Final implementation of prototype.

Late May: Finalization of project.

List of Key References:

References

Anon., n.d. *The KITTI Vision Benchmark Suite.* [Online]
Available at: http://www.cvlibs.net/datasets/kitti/index.php

This section is to be filled in by the representative of the Institute Research Sub-Committee prior to forwarding of this Form to the 'MCAST Research Ethics Committee' for final ethics approval:

Nature of ethical consideration	***Outcome (*)***	***Comments***
Consideration of possibility of issues regarding confidential personal data:		
Consideration of possibility of physical harm		
Consideration of possibility of moral harm		
Consideration of possibility of business harm		

(*) Legend to record outcome by Institute Research Sub Committee:

A *- Ethical considerations have been* ***addressed appropriately*** *by Researcher;*
B *- No (****Nil****) relevant ethical considerations are applicable under purpose of study as described by Researcher.*
C *- Ethical consideration have* ***not been addressed appropriately*** *by Researcher;*
D *- Applicable ethical consideration have* ***not been considered*** *by Researcher.*

Details of Representative to the 'Institute Research Sub-Committee.	
Name	Signature
Designation	Date