Faculdade de Engenharia da Universidade do Porto

Visual-Inertial based autonomous navigation of an Unmanned Aerial Vehicle in GPS-Denied environments

Francisco de Babo Martins

EEC0035 - Preparação da Dissertação

U. PORTO

FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Mestrado Integrado em Engenharia Electrotécnica e de Computadores

Supervisor: Luís Teixeira

February 18, 2015

“Roads?
To where we are going, we don’t need...roads!”

Dr. Emmett Lathrop Brown

Contents

List of Figures

Abbreviations and Symbols

UAV	*Unmanned Aerial Vehicle*
MAV	*Micro Air Vehicle*
GPS	*Global Positioning System*
IMU	*Inertial Measurement Unit*
SURF	*Speeded-Up Robust Features*
SVM	*Support Vector Machines*
HOG	*Histogram of oriented gradients*
SIFT	*Scale-invariant feature transform*
SLAM	*Simultaneous Localization and Mapping*
SVM	*Support Vector Machines*
HOG	*Histograms of Oriented Gradients*

Chapter 1

PDI report introduction

The aim of this document is that of presenting all the work developed throughout the semester for the curricular unit *Preparation for the MSc Dissertation*.

This report is organized as follows.

In Chapter 2, the problems to be approached during this project will be presented and described.

Chapter 3 will present the state of the art of the relevant topics regarding the project development and a brief description of the methods to be profoundly researched.
The aforementioned methods consist of:

- Object detection.
- Object tracking.
- Machine Learning.

Conclusively, chapter 4 contains the work plan for the semester in which this project will take place. Additionally, a description of the technologies to be used will be described as well.

1.1 Project objectives and motivations

The aim of this project is to allow an UAV to safely navigate in a previously unknown and GPS-denied environment without requiring pre-made maps or precoded information to avoid collisions. Using on-board cameras, which are indispensable components of an UAV, the autonomous navigation will be achieved by detecting and tracking visual features and objects in order to gift an UAV of a collision free navigation. In order to reduce the on-board payload data processing, the UAV will send the acquired data to an external processing unit (laptop). Autonomous indoor navigation can be extremely useful in search and rescue operations[1], surveillance tasks of areas hazardous to humans, traversal of paths inaccessible to ground vehicles and quick deployment.

In order to achieve this goal, the system needs to meet the following requirements:

- The computational complexity of the system should be independent of the environment size.

- Firm and robust feature detection and object tracking.
- Fast data processing to satisfy real time constrains.
- Brisk obstacle avoidance;
- Flexible navigation;

The feature detection and object tracking to be performed is of the utmost importance as the very own navigation and obstacle avoidance depend on it.

Chapter 2

Problem statement

2.1 Defining the problem

Being able to hover, fly laterally and at low speeds, makes UAVs an optimal platform to accomplish different military and civilian tasks like reconnaissance support in hazardous zones, visual surveillance and inspection. In addition, some relevant industries are starting to use drones for other tasks beyond surveillance (p.e Amazon's "Prime Air"). Moreover, the most important task in order to achieve UAV autonomy is autonomous navigation. This may prove useful in a near future for tasks in indoor environments such as indoor transportation, object retrieval (p.e a missing part in a assembly line), monitoring misplaced books in a library and autonomously reporting sports events.

Although years of research of GPS position and data tracking have improved outdoor navigation and localization, in environments such as indoors or dense urban areas where maps are unavailable and the GPS signal is weak, a UAV will operate in high hazardous regions, running the risk of becoming lost and colliding with obstacles.

Since the gist of this project consists of enabling an UAV to autonomously navigate in a unknown environment without resorting to GPS localization, the main problem is using visual odometry and on-board IMU to develop localization, estimation and planning algorithms to achieve a autonomous collision and obstacle avoidance navigation. A vast majority of UAVs depend on GPS for navigation, hence this project is more challenging since the GPS is denied.

This project is therefore focused on exploring methods that may allow autonomous flight indoors without GPS and previous knowledge about the environment while using only on-board sensors.

2.2 Refining the problem

Given the complexity of the stated problem, this project needs to be split into simpler and implicit problems:

- **Acquisition and quantification of sensory data:** Processing and evaluating the values from the UAV sensors is of the utmost importance. The readings from the sensors will prove to be a valuable asset when the development of the autonomous navigation algorithm requires sensory data to analyse the state of the UAV.
- **Adjustment of the UAV navigation parameters:** In order to create a navigation algorithm, a fully functional communication between the UAVs and the external processing unit (laptop) must be established. Prior to the creation of the navigation algorithm, it must ensured that the laptop is fully capable of sending desirable commands to the UAV. These include adjustments to parameters like yaw, pitch and roll.
- **Image data acquirement from the UAV camera:** All the image frames captured by the UAV camera will be used to compute all the necessary data for the navigation algorithm.
- **Image processing:** Images previously acquired should undergo a series of processes aimed at detecting points and regions of interest which are relevant to the UAV obstacle avoidance navigation. Such processes may include: Binarization, Segmentation, Noise Reduction and Perspective Transformations.
- **Image feature detection and extraction:** All the acquired frames from the camera should be analysed in order to detect features of interest to be used on the classification stage. The techniques to be used may include: Canny edge detector, SURF [2], HOG, and Hough transform.
- **Feature matching and classification**: The obtained features will be used to train a classification model to be used by the UAV to detect safe and hazardous zones while flying autonomously. This may be achieve using: Bag-Of-Words, SVM for pattern matching.

Chapter 3

Literature Review

3.1 Introduction

This chapter will present a literature review which is befitting for a better understanding to the problem under studied. In section 3.2 will be presented a brief description of the current development state of autonomous navigation using UAVs. Section 3.3 will focus on some currently used algorithms for image processing, machine learning and navigation that may prove to be the most suitable methods for the problem at hand.

3.2 UAV Autonomous Navigation

In recent years, UAVs such as quad-copters have increasingly gained interest in robotics and computer vision research. Commonly used as a flying camera, they are mostly employed for surveillance tasks and recording humanly impossible videos. Therefore, in order to navigate safely, these equipments rely on their on-board sensors and cameras to track their position autonomously. The advantages of using UAVs includes their easy manoeuvrability and ability to hover.

Whilst there has been substantial research about navigation algorithms on ground vehicles, these approaches can not be directly transferred to aerial vehicles:

- An indoor aerial vehicle cannot hold as many sensors as a ground vehicle.
- A flying vehicle as more degrees of freedom than a ground vehicle, which prevents the use of 2D navigation algorithms.
- The risk of failure is more catastrophic in a flying vehicle. If a navigation prediction failure occurs on a ground vehicle, it can be stopped by a user, whereas when the same occurs in a flying vehicle, it falls on the ground damaging its structural integrity.

It is of the utmost importance to gift an UAV with collision avoidance capabilities to avoid the risk of damaging the structure of the vehicle and the surrounding environment. This catastrophic scenario depends mostly on the altitude of the vehicle and the nature of the collision.

3.2.1 On-board Electronics

Different UAVs have an array of devices capable of providing accurate data which can prove to be useful for navigation, localization, landing and obstacle avoidance tasks.

3.2.1.1 GPS

The Global Positioning System [3] was funded by the United States Department of Defence and was initially designed for the United States military, although in 1980, it was made available for civilian use.

This system works anywhere in the world, 24 hours a day and in any weather conditions. Four GPS satellite signals are used in order to compute the positions in three dimension and also the time offset in the receiver clock.

3.2.1.2 IMU

The IMU is a single unit, that houses two sensors, which collects angular velocity and linear acceleration data and returns it to the main processor.It is used to acquire pitch, roll and yaw data from the UAV.

A brief description of these two sensors follows:

- **Accelerometer** An accelerometer is a device that measures acceleration forces which may be static like the force of gravity, or dynamic like vibrating or even moving the accelerometer itself.

- **Gyroscope** A Gyroscope is a device that measures the orientation of a device based on the angular momentum of that same device. It is used to acquire the angular rate of a certain vehicle[4].

3.2.1.3 Laser Range Finder

This device uses a laser beam to determine the distance of the source of emission to an object. Its normal mode of operation consists of sending a laser pulse to a certain object and measuring the distance that it takes to the pulse to reflect off the object and bounce back to the emitter. They are commonly used for 3D-Modelling[5].

3.2.1.4 Ultrasonic Sensor

This sensor can be used as a sonar based altimeter, providing altitude estimation and vertical displacements. Their modus operandi is similar to a sonar in which consist of generating high frequency sound waves and measuring the time that takes for the transmitter to received the echo of the emitted wave, thus determining the distance to an object.

3.2.1.5 Pressure Sensor

A pressure sensor is commonly used to provide stability to an UAV, allowing the on-board processing to automatically correct and maintain a still position while the vehicle is airborne. This is achieved regardless of altitude and wind intensity[1] pressure sensors on-board provide unique stability that will automatically correct and maintain a still position in the air regardless of altitude and wind up to

3.2.1.6 Magnetometer

A Magnetometer is a electronic device capable of measuring the strength and direction of a magnetic field. Therefore it can be used in navigation applications by measuring the earth's magnetic field.

3.2.1.7 Monocular Camera

Even though this type of cameras have low weight, power consumption, size and cost, they still provide an immense amount of information, which is unmatched by any other type o sensors. Compared to other depth measuring devices, the range of a monocular camera is virtually unlimited. This feature proves to be useful when using a monocular SLAM system to operate in open environments[6].

In spite of those advantages, this device proves to be inefficient when it comes down to determine the scale of the environment as there is no way to obtain depth perception. In this case, a vehicle must rely on another set of sensors like those of the IMU (3.2.1.2).

3.2.1.8 Stereo Camera

This particular camera, containing two or more lenses with a image sensor per lent, is a direct solution to the disadvantage of the aforementioned monocular camera: extraction of depth information of an object in a image. It is possible to capture 3D images with these cameras due to the fact that the lenses simulate the human binocular vision. Hence, this process is called *stereo photography*.

Not all stereo cameras are used to acquire 3D photographs or video whereas depending on the lenses configuration, the acquired data may not contain stereoscopic information.

3.2.1.9 RGB-D Sensor

The RGB-D sensor provides the combined data of RGB color information and per-pixel depth information. The data acquired from these sensors can be represented as a *point cloud* which consists of a collection of points in three dimensional space. Additional features can be associated with the point from that space, meaning that color can be one of those features.

One device that provides these functionalities and is affordable, unlike the rest of the RGB-D sensors in the market, is the Microsoft Kinect[7].

3.2.2 Parrot AR-Drone

The Parrot AR-Drone is an electrically powered quad-copter who was originally intended for augmented reality games. Its sensory equipment consists of:

- A three-axis accelerometer and gyroscope.
- A Sonar-based altimeter.
- Two cameras:
  - The frontal camera has a 75x60 field of view and provides a 320x240 pixel color image. When object detection is the focus of a project, it is carried out by this camera using computer vision algorithms.
  - The vertical camera has a 45x35 field of view and provides a 176x144 pixel color image.

Figure 3.1: AR-Drone with the indoor hull

Since it was designed as a toy, this drone has limited computational power which may prove a problem as image processing using computer vision algorithms require a lot of resources. One possible workaround is using a computer that is connected to the AR-Drone via wireless. The communication flow between the AR-Drone and the computer can be seen in figure 3.2.

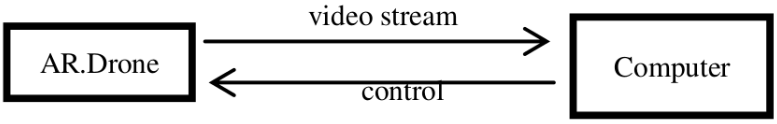

Figure 3.2: Data exchange between the AR.Drone and a Computer

The batteries of this UAV provide enough energy for a 15 minutes flight. However, its turbines are not powerful enough to keep a steady position in windy conditions, which limits its outdoor use.

3.2.2.1 AR-Drone Brown Package

The SDK provided by the AR-Drone manufacturer does not provide the desirable tools to use in navigation projects. Therefore, in order to have a better control over the AR-Drone, the **AR-Drone Brown Package** was developed[8].

This package is used to connect a computer to the AR-Drone and consists of the SDK from Parrot with a ROS wrapper.

The Drone *teleop* package is used to issue commands to the AR-Drone. The commands that can be sent are:

- Take off/land.
- Forward, Backward, Left and right.
- Rotate left and right.
- Increase/decrease altitude.
- Toggle emergency state.

3.2.3 Research and Development

The goal of developing autonomous UAVs is one of allowing them to perform complex manoeuvres and navigate independently in structured and unstructured environments.

3.2.3.1 Laser Assisted Navigation

A relevant project was that of a team from the Massachusetts Institute of Technology which used a laser range-finder sensor to enable an MAV to autonomously explore and map unstructured and unknown environments[9]. They used on-board sensors to estimate the position of the MAV and used SLAM algorithms to build a map of the environment around the vehicle.

The motion estimation of this project was comprised of two different problems: motion estimation using laser scan-matching and height estimation using the on-board IMU.

3.2.3.2 Sonar Assisted Navigation

Sonar technology used to be prevalent among UAVs and some projects still used it for field mapping using UAVs [10]. However, due to the fact that the time it takes for the sonar waves to travel decreases the UAV obstacle detection range, this technology can no longer be used.

The alternative is a way of detection images and avoid obstacles much sooner which can be achieve using a technology that does not depend on propagation waves, but rather uses light as medium to detect objects: optical cameras.

3.2.3.3 Camera Assisted Navigation

Pre-coded Maps

A team from the Czech Technical University in Prague developed a navigation system for a UAV using a monocular camera and odometry[11]. The navigation system was based on dead-reckoning techniques[12] to calculate the traversed segment length and the information from the camera to calculate the UAV yaw and vertical speed. This allowed the team to use a histogram voting scheme.

The developed navigation method consists of a "record and play" technique in which the UAVs is guided along the path that it will autonomously navigate later on. During the initial phase, a map is created by guiding the UAV along a path, using the front-facing camera to recognize and track salient features using the SURF[2] feature detector. In the second phase the UAVs navigates autonomously in the learned path by using the dead-reckoning system to calculate the traversed distance and retrieves the relevant landmarks from the precoded map in order to compare to the data currently being obtained through the camera. The difference of coordinates is then computed using a histogram voting technique and this information is then processed to the UAV yaw and altitude controllers to safely adjust the drone position. Briefly, using this method, the UAV builds a SURF-based landmark map in a guided tour and then uses the method described above to navigate autonomously in the mapped environment.

The only limitation of this project consists in the necessity of using a pre-autonomous-navigation phase in which the UAV is human-guided.

Autonomous Indoor Navigation

The project to be implemented throughout this semester is the spiritual successor of a project of a team from University of Texas [13] which aimed at navigating autonomously in indoor environments (corridors) and industrial environments (production lines) and detecting and avoiding obstacles (people). The navigation was accomplished by using the vanishing point algorithm, the Hough transform (3.3.1.5) for wall detection and avoidance and HOG descriptors using SVM classifiers for detecting pedestrians. The UAV used was the AR-Drone.

Firstly, edge detection is performed using the Canny edge detector followed by a Gaussian blurring to reduce noise and irrelevant edges. Using the points exported by the Canny edge detector, the Hough transform is used to detect the perspective lines using a voting scheme procedure which finds the lines that are closer to those same points. The following operation was that of removing some lines that were dependent on thresholds that influenced the Canny edge detector and the Hough transform (outliers). This was achieved by thresholding their angles (in radians).

The next phase was finding the line intersection of the lines which were detected by the Hough transform, which was achieved by using the Least square problem which estimates the minimum distance of a point from a set of lines. In corridors, like those used in this project, the line interception is called the vanishing point. In this step there is a boundary box in the image used to distinguish between good estimations of the vanishing point and bad ones. If there is a bad estimation, it is overlook and the system awaits for another one.

As the AR-Drone lacks of a laser or even an obstacle avoidance sensor, only the camera can be used to avoid collisions with the right and left wall. In this project it was assumed that the UAV should have the ability to know its relative position to the left and the right walls in order to avoid them while moving. The way they did it was by manually pre-estimating this ability which consisted of manually putting the camera of the drone to the right and left wall and analysing the angular value of the lines obtained from the Hough transform. It is important to mention that the angular values of the lines can be automatically estimated while the drone is at the center of a corridor. By exporting the value of the angles show in figure 3.3 it was possible to determine when there was enough space for the UAV to safely navigate without hitting the side walls:

- When angle angle $B = D$, the drone is positioned at the center of the corridor.
- When the angle A is twice bigger than the the angle $B + D$, the AR-Drone has enough space to navigate without hitting any obstacles on the side walls or even the side walls themselves.

Figure 3.3: Hough Lines in the left and right wall

The human detection process implemented for this project used the Histogram of Oriented Gradient descriptors (HOG) and a linear SVM classifier (3.3.2.1). The SVM classifier was used to test the feature detector in the MIT pedestrian database which contains 1800 pedestrians images of different poses and backgrounds.

In order to avoid hitting the target (person), the boundaries of the side walls are considered a place of safe passage, meaning that when a person is detected by the AR-Drone, an estimation of the most large free sideways are is made in order to avoid hitting the moving target. The wall detection algorithm is of the utmost importance as the position of the moving target is computed relatively to the boundaries of the side walls.

Stereo and SLAM

Other approaches involved using stereo cameras[14] and the Microsoft Kinect RGB-D sensors[15] to solve the SLAM problem and find the best solution for a robust state estimation and control methods. These approaches perform local estimation of the vehicle position and build a 3D model of the environment (SLAM) to plan trajectories through an environment.

3.2.3.4 IMU Assisted Navigation

Some other projects recurred to Kalman filters and ultrasound sensors to compute the UAV pose estimation in a GPS denied environment[16]. Every time that the UAV has to reach a certain location, a path which provides a good observation density is computed from a predefined map, minimizing the risk of localization failures.

3.3 Methods and Algorithms

3.3.1 Image Processing

Image processing is a field of imaging science in which the output of image processing results in yet other image or a set of parameters and characteristics related to the image.

Commonly, the first step after capturing and image is that of converting it into a greyscale image and then into a binary image, or changing its color space from RGB to HSV.

Concerning the the identification of connected components, a segmentation and contour detection is used and in order to reduce computational costs, objects smaller than a specific are are normally discarded.

In order to detect shapes and objects in the video feed of an UAV the methods presented in the following sections are to be considered.

3.3.1.1 Object Detection

The gist of object detection is the precise recognition of an object in a set of images. This is a very complex problem and nowadays still proves to be quite challenging[17]. Although the most used approach to detect objects in a video stream consists of using information from different frames, it delivers results with a high error rate. A typical workaround is that of analysing temporary information which was computed from a set of sequential frames, hence reducing the error rate[17].

3.3.1.2 Edge Detection

In image processing, an edge is defined as a discontinuity in rightness of an image. Therefore the problem of detecting edges consists of finding those discontinuities which may represent changes in properties of that same image like depth, orientation and illumination.

The *Canny edge detector* is one of the most used edge detection operators and is comprised of the following steps:

- Noise removal with Gaussian filer.
- Localization of the intensity gradients of the image.
- Non-maximum suppression application to eliminate false responses to edge detection.
- Double threshold application to resolve potential edges.
- Suppress the remaining edges that are week and are not connect to strong edges.

3.3.1.3 Feature Detection

The SURF algorithm[2] is a robust local feature detector built upon the SIFT descriptor but is reported to perform better than the latter[18] when it comes down to speed and robustness to viewpoint and illumination changes.

This algorithm detects salient features and provides their image coordinates together with their descriptions. It is composed of 4 stages:

- Integral image generation
- Interest point detection
- Descriptor orientation assignment
- Descriptor generation

In a navigation algorithm the calculation of the Speeded up Robust Features may prove to be the most intensive and computationally demanding part, as it needs to be accelerated in order to obtain a real-time performance.

3.3.1.4 Object Tracking

Similar to object detection, object tracking consists in the process of tracking an object in a sequence of frames or images. Its typical applications are in surveillance tasks and interaction systems.

Some typical challenges that object tracking tries to overcome and which directly influences its success are the following:

- The movement of the object.
- A change in the pattern, structure and background of the object.
- Camera movement.
- Object occlusion by other object or background.

The major algorithms used in object tracking are Point Tracking, Kernel Tracking and Silhouette Tracking.

Optical Flow

When it comes down to object tracking the concept of optical flow is often used. Being the apparent motion between an observer and de environment, it is used to estimate the motion field. This is accomplished by estimating the location of each pixel from one frame in the second frame. However it does not inescapably corresponds to actual motion.

One of the major challenges in calculating optical flow is that of finding corresponding pixels as many of them can have the same color and even surrounding pixels with the same color value.

This leads to the assumption that in order to calculate the optical flow, it is assumed that there is very little movement from one image to the other, leading to a "travelled distance" of just one pixel. This assumption poses a problem when dealing with fast moving vehicles due to delays in sending and interpretation images.

One possible workaround is that of using sparse optical flow. This is a process in which image features are tracked across different images[19].

Some navigation methods used on UAVs are based on a dead-reckoning system, which is the process of calculation a current position based on a previously determined one. These methods are used to estimate the speed of an aerial vehicle based on the optical flow in the image of its bottom camera. The value of these speeds are then integrated in order to get an estimate of the drone position[1].

3.3.1.5 Feature Extraction

Hough Transform

One of the most used feature extraction techniques is the Hough Transform. It was originally designed to find lines in a image[20] but has been modified in order to detect shapes like circles and ellipses.

Many variants of the standard Hough Transform were developed, such as the Gerig Hough Transform with Gradient, 2-1 Hough Transform and Fast Hough Transform.

Prior to shape detection, an image must undergo a preprocessing in order to acquire information of the edge of each object. Techniques like those mentioned in 3.3.1.2 are used to compute this data which will later be processed using the Hough Transform to find a shape.

In order to find a line, which is represented by

$$y = mx + c \tag{3.1}$$

the first thing to do is to represent a line into a point, resulting in the following equation:

$$c = -mx + y \tag{3.2}$$

The Hough Transform converts a point in xy-space to a line in mc-space (the parameter space), which consists of representing the information of an image discrete points $(x_1, y_1)(x_2, y_2)$ in terms of the slope parameter m and the intercept parameter b. As the points from an edge of an image are therefore transformed into lines, some lines that are constructed from edges of the same object will intersect. This means that the intersection of lines in in mc-space correspond to information of a line in xy-space.

However, a vertical lines proves as problem as m has an infinite value, requiring infinite memory to store the data in mc-space. Therefore, another representation is used to avoid this problem.

In this new representation, a line is represented by the distance p of a line from the point (0,0) and the angle θ of a line to the x-axis. Thus, a line is represented by the following equation:

$$p = x.cos(\theta) + y.sin(\theta) \tag{3.3}$$

Where x and y is a point passed by the line.

In this method, a line in xy-space will be transformed into a point in $p\theta$-space and a point in xy-space will be transformed into a sinusoidal curve[21].

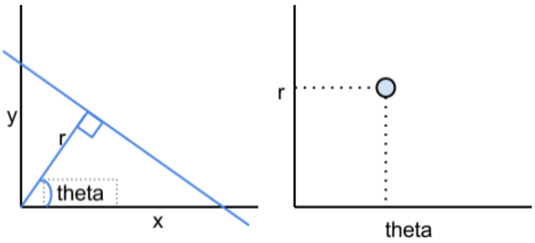

Figure 3.4: Representation of a line

Using this method, the Hough Transform uses the values of x, y and R to recognize circles in a image. The Hough Transform will create circles with a certain radius in every pixel of an object and the calculation of the points with the maximum number of interceptions of those circles will be computed. In the end, the point with the most interceptions will acknowledge as the center point.

The Hough transform has proven to be a strong candidate for this project as it has proven to be efficient and quite reliable in some other similar projects[22].

A team from the university of Indonesia used the Hough Transform method[23] to detect and track shapes in an indoor environment using the front-facing camera of the AR-Drone. They accomplish this by following these steps:

- Conversion of the image color model from RGB to HSV.
- Conversion of the video stream into a binary image: since this team detected objects based on its color and shape, they had to pay close attention to the range of the HSV components in order to obtain the ideal threshold.
- Following, using a implementation of the Hough Transform method from the OpenCV library, a circle shape is found.
- The tracking method then uses the information obtained in the previous step and sets a rectangle window inside the viewfinder of the input image. After, the controller method uses this information and makes sure that the rectangle always contains the information of the desirable object to track.

3.3.2 Machine Learning

Machine learning is a sub-field of artificial intelligence that enables computers to act independently without being programmed to do so. Many applications that rely on machine learning are,for example, self-driving cars, practical speech recognition and effective web search. This is accomplished by constructing algorithms that are able to learn from available data.

There are different methods that can be categorized as follows:

- Supervised learning which use both the inputs and the outputs in order to replicate an intended model:
    - Parametric algorithms.
    - Non-Parametric algorithms.
    - Support Vector Machines (SVM).
    - Kernels.
    - Neural networks.
- Unsupervised learning, which are used when there is no information regarding the intended output:
    - Clustering.
    - Dimensionality Reduction.
    - Recommender Systems.
    - Deep learning.

3.3.2.1 Support Vector Machines

In order to gift an UAV with the ability to recognize objects, the machine learning method that is considered to be of relevant use is Support Vector Machines (SVM).

SVMs are a kernel-based technique that aim to output an optimal separating hyperplane that categorizes previously labelled images. They are mostly used in patter recognition tasks and require a training set of images in order to train the classifier. Its modus operandi is finding the hyperplane that maximizes the margin of the training data. This means that the optimal separating hyperplane gives the largest minimum distance to the training data.

SVMs have a fairly good success rate towards redundant attributes, have a reliable accuracy and classify data rather quickly. All this makes SVMs a good choice for this project.

Chapter 4

Work Plan

In order to achieve a successful result in any project, there is the need to always plan ahead. This allows to prioritise objectives and allocate resources accordingly.

A timeline containing the tasks and objectives to be achieved throughout this project is presented in section 4.1. A brief description of the project stages is also provided. In section 4.2 a description of the methodology to be used is presented. Section 4.3 will present the technologies and tools that will aid the development of this project. To conclude, in section 4.4, the tasks that have already been completed are presented.

Its important to bear in mind that the presented calendarization is merely a proposal as the different tasks and the time associated with each of them may be shifted throughout the duration of the this project.

4.1 Calendarization

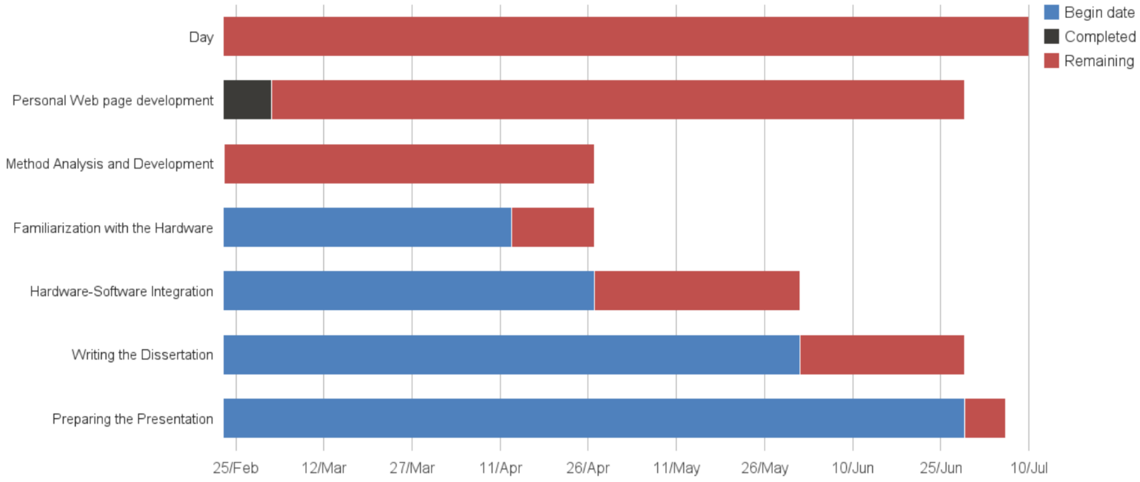

Figure 4.1: Gantt Diagram

Table 4.1 presents the sets of the project and the corresponding duration. Bellow, a brief description of each set is also provided.

Table 4.1: Project Tasks Scheduling

Task	Initial Date	Duration(weeks)	Ending Date
Personal Web page development	18-02-2015	18	29-06-2015
Method Analysis and Development	23-02-2015	9	27-04-2015
Familiarization with the Hardware	13-04-2015	2	27-04-2015
Hardware-Software Integration	27-04-2015	5	1-06-2015
Writing the Dissertation	1-06-2015	4	29-06-2015
Preparing the Presentation	29-06-2015	1	06-07-2015

- **Personal Web page development:** Throughout the development of the project certain documents like reports from the meetings with the advisor and weekly reports will need to be presented. Therefore, a website containing all information will be created and constantly updated.

- **Method Analysis and Development:** From all the image processing and machine learning methods available, those who prove to be useful for the development of this project will be analysed and use for development.

- **Familiarization with the Hardware:** Once the development has reach a phase were it can be tested on an UAV, test regarding the integration with the hardware of the Parrot A.R Drone will be performed.

- **Hardware-Software Integration:** The algorithms developed until this point will be adapted to the Parrot A.R Drone in order to evaluate the real time performance of the system and take action if the need for some modifications to the project arise.

- **Writing the Dissertation:** The last part of the project will be the redaction of a report in which all the research and results will be compiled.

- **Preparing the Presentation:** A presentation of the project will also be prepared in order to demonstrate all the work done along the duration of this project.

4.2 Methodology

Throughout the duration of this project, the following methodology will be followed:

- Acquire the required hardware for this project (Parrot AR-Drone).

- Research the image processing methods most suitable for the task at hand.

- Research the machine learning methods that may be most useful for categorizing features in the drone video stream.

- Decide which approach to follow regarding and which methods to use.
- Divide the system into three phases (specify and design the architecture of the system), and implement each one using the previously researched methods:
    - Autonomous flight.
    - Object recognition.
    - Avoidance commands.
- Redaction of the final report and presentation.

4.3 Technologies, tools and work platforms

All the documentation will be done in LaTeXand will be available on the personal website that will be created for that matter.

The implementation of the methods and algorithms to be used in this project will be done using the *C++* programming language. The OpenCV[24] library will be used to ensure fast and efficient image processing.

The UAV chosen for this project is the Parrot AR.Drone[1] which is complemented with advanced features and the development from the software side is relatively easy compared to other quad-copters.

4.4 Concluded tasks

Throughout the duration of curricular unit Preparation for the MSc Dissertation the following milestones were achieved:

- Preliminary research of the subject of the project.
- Understanding of the different approaches to object detection and tracking.
- Study of some methods such as the Hough Transform and SVMs.
- Research about the AR-Drone hardware and software
- Understanding of navigation nomenclature
- Familiarization with the OpenCV library for C++
- Familiarization with LaTeX

References

[1] Pierre-Jean Bristeau, François Callou, David Vissière, and Nicolas Petit. The Navigation and Control technology inside the AR . Drone micro UAV. *Proceedings of the 18th IFAC World Congress, 2011*, 18:1477–1484, 2011.

[2] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. Surf: Speeded up robust features. In *Computer vision–ECCV 2006*, pages 404–417. Springer, 2006.

[3] Navigation Center general information on gps,
http://www.navcen.uscg.gov/?pageName=gpsmain, 2015-02-09.

[4] Gyroscopes uses for gyroscopes
http://www.gyroscopes.org/, 2015-02-09.

[5] Hartmut Surmann, Kai Lingemann, Andreas Nüchter, and Joachim Hertzberg. A 3d laser range finder for autonomous mobile robots. In *Proceedings of the 32nd ISR (International Symposium on Robotics)*, volume 19, pages 153–158, 2001.

[6] Jakob Engel, Jürgen Sturm, and Daniel Cremers. Camera-based navigation of a low-cost quadrocopter. In *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, pages 2815–2821. IEEE, 2012.

[7] Kinect for windows, *http://www.microsoft.com/en-us/kinectforwindows/*, 2015-02-13.

[8] Ros driver for the parrot ar.drone,
https://github.com/AutonomyLab/ardrone_autonomy, 2014-12-15.

[9] Abraham Bachrach, Samuel Prentice, Ruijie He, and Nicholas Roy. Range–robust autonomous navigation in gps-denied environments. *Journal of Field Robotics*, 28(5):644–666, 2011.

[10] Mohamad Farid bin Misnan, Norhashim Mohd Arshad, and Noorfadzli Abd Razak. Construction sonar sensor model of low altitude field mapping sensors for application on a uav. In *Signal Processing and its Applications (CSPA), 2012 IEEE 8th International Colloquium on*, pages 446–450. IEEE, 2012.

[11] Tomáš Krajník, Matías Nitsche, Sol Pedre, Libor Přeučil, and Marta E. Mejail. A simple visual navigation system for an UAV. In *International Multi-Conference on Systems, Signals and Devices, SSD 2012 - Summary Proceedings*, 2012.

[12] Ariane S Etienne, Roland Maurer, Josephine Georgakopoulos, and Andrea Griffin. Dead reckoning (path integration), landmarks, and representation of space in a comparative perspective. 1999.

[13] Alexandros Lioulemes, Georgios Galatas, Vangelis Metsis, Gian Luca Mariottini, and Fillia Makedon. Safety challenges in using ar. drone to collaborate with humans in indoor environments. In *Proceedings of the 7th International Conference on PErvasive Technologies Related to Assistive Environments*, page 33. ACM, 2014.

[14] Markus Achtelik, Abraham Bachrach, Ruijie He, Samuel Prentice, and Nicholas Roy. Stereo vision and laser odometry for autonomous helicopters in gps-denied indoor environments. In *SPIE Defense, Security, and Sensing*, pages 733219–733219. International Society for Optics and Photonics, 2009.

[15] Albert S Huang, Abraham Bachrach, Peter Henry, Michael Krainin, Daniel Maturana, Dieter Fox, and Nicholas Roy. Visual odometry and mapping for autonomous flight using an rgb-d camera. In *International Symposium on Robotics Research (ISRR)*, pages 1–16, 2011.

[16] James F Roberts, Timothy Stirling, Jean-Christophe Zufferey, and Dario Floreano. Quadrotor using minimal sensing for autonomous indoor flight. In *European Micro Air Vehicle Conference and Flight Competition (EMAV2007)*, number LIS-CONF-2007-006, 2007.

[17] Liming Wang, Jianbo Shi, Gang Song, and I-fan Shen. Object detection combining recognition and segmentation. In *Computer Vision–ACCV 2007*, pages 189–199. Springer, 2007.

[18] David G Lowe. Object recognition from local scale-invariant features. In *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, volume 2, pages 1150–1157. Ieee, 1999.

[19] Jianbo Shi and Carlo Tomasi. Good features to track. In *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference on*, pages 593–600. IEEE, 1994.

[20] Gary Bradski and Adrian Kaehler. *Learning OpenCV: Computer vision with the OpenCV library*. " O'Reilly Media, Inc.", 2008.

[21] John Illingworth and Josef Kittler. A survey of the hough transform. *Computer vision, graphics, and image processing*, 44(1):87–116, 1988.

[22] Cooper Bills, Joyce Chen, and Ashutosh Saxena. Autonomous mav flight in indoor environments using single image perspective cues. In *Robotics and automation (ICRA), 2011 IEEE international conference on*, pages 5776–5783. IEEE, 2011.

[23] Futuhal Arifin, Ricky Arifandi Daniel, and Didit Widiyanto. Autonomous detection and tracking of an object autonomously using ar. drone quadcopter. *Jurnal Ilmu Komputer dan Informasi*, 7(1):11–17, 2014.

[24] Opencv, *http://opencv.org/*, 2015-02-10.