AUTOMATED FISH DETECTION AND IDENTIFICATION

WONG POH LEE

UNIVERSITI SAINS MALAYSIA

2015

AUTOMATED FISH DETECTION AND IDENTIFICATION

By

WONG POH LEE

Thesis submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy

October 2015

ACKNOWLEDGEMENT

First of all I would like to express my gratitude to my supervisor, Prof. Abdullah Zawawi Talib (Supervisor in Universiti Sains Malaysia) for the useful guidance, supervision and feedbacks of this thesis. I would also like to thank Prof. Jean-Marc Ogier and Prof. Jean-Christophe Burie (Supervisor and Co-Supervisor in University of La Rochelle respectively) for their guidance and supervision in making this research a success. They have been giving great feedback and comments in making this research possible. In addition, I would like to thank my co-supervisors, Mohd Azam Osman and Assc. Prof. Dr. Khairun Yahya for giving great suggestions, overcoming difficulties and supporting me throughout the process of the research.

Special thanks also go to José Menneson from Lille University of Science and Technology for his help and providing great solutions in solving problems. He has given a lot of great suggestions and guidance in overcoming object characterisation issues which are mainly used in the area of image processing and computer vision.

Subsequently, I owe more than thanks to my family members which include my mother and my brother for their continuous support and encouragement throughout my life. It is impossible for me to achieve and graduate seamlessly without their support.

During my period in this research, my friends are helpful in giving advises and knowledge for me to improve and to make this a possibility in achieving my goals in pursuing up to this level. I would like to acknowledge all my friends for their support.

Finally, I would like to acknowledge the contribution of the School of Computer Sciences, USM and the RU Grant (1001/PKOMP/817070) from USM. I would also like to thank the MFUC (Malaysia France University Centre) for their support in the cotutelle programme between Universiti Sains Malaysia and University of La Rochelle.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

ADV	Acoustical Doppler Velocimetry
ANN	Artificial Neural Networks
BRIEF	Binary Robust Independent Elementary Features
BRISK	Binary Robust Invariant Scalable Keypoints
COIL-100	Columbia Object Image Library 100
EMD	Earth Mover’s Distance
FAST	Features from Accelerated Segment Test
FERET	Facial Recognition Technology
GCFD	Generalized Colour Fourier Descriptor
$GCFD_{adapt}$	Adapted Generalized Colour Fourier Descriptor
GMM	Gaussian Mixed Model
HSTA	Hierarchical Spatiotemporal Data Association
HSV	Hue, Saturation, Value
KLT	Kanade-Lucas-Tomasi
MD	Motion Detection
MDPF	Motion Detection with Particle Filter
$MDPF_{cache}$	Motion Detection with Particle Filter
OpenCV	Open Source Computer Vision
PF	Particle Filter
PF_{cache}	Particle Filter with Added Cache
RGB	Red, Green, Blue
SFU	Simon Fraser University
SIFT	Scale Invariant Feature Transform
SURF	Speeded Up Robust Features
SVM	Support Vector Machine
USB	Universal Serial Bus

LIST OF PUBLICATIONS

Wong, P. L., Osman, M. A., Talib, A.Z., & Yahya, K. (2012). Modelling of Fish Swimming Patterns Using an Enhanced Object Tracking Algorithm. *Frontiers in Computer Education, Advances in Intelligent and Soft Computing, 133(1), 585-592.*

Wong, P. L., Osman, M. A., Talib, A.Z., & Yahya, K. (2013). Tracking Multiple Fishes Using Colour Changes Identification and Enhanced Object Tracking Algorithm. *Applied Mechanics and Materials, ISSN: 1662-7482, 284-287(1), 1528-1532.*

Wong, P. L., Osman, M. A., Talib, A.Z., Ogier, J., & Yahya, K. (2014). Tracking Multiple Fish in a Single Tank using an Improved Particle Filter. *Lecture Notes in Electrical Engineering, 279(1), 799-804.*

Wong, P. L., Osman, M. A., Talib, A.Z., Burie, J. C., Ogier, J., Yahya, K., & Mennesson, J. (2015). Recognition of Fish Based on Generalized Color Fourier Descriptor. *Proceedings of the Science and Information Conference 2015, London, United Kingdom, July 28-30, 2015, 680-686.*

Wong, P. L., Osman, M. A., Talib, A.Z., Burie, J. C., Ogier, J. & Yahya, K. (2015). Enhanced Object Tracking in Real-time Environment using Dual Camera. *4th International Visual Informatics Conference 2015 (IVIC'15) Bangi, Malaysia, November 17-19, 2015. (Accepted)*

PENGESANAN DAN PENGENALPASTIAN IKAN BERAUTOMATIK

ABSTRAK

Pengecaman dan pengenalpastian ikan menggunakan kaedah komputan telah menjadi sebuah bidang penyelidikan yang amat popular dalam kalangan para penyelidik. Kaedah-kaedah berkenaan penting kerana maklumat yang dipaparkan oleh ikan seperti pola trajektori, lokasi dan warna boleh menentukan sama ada ikan berkenaan sihat atau sakit. Pada masa ini, kaedah-kaedah ini tidak begitu tepat terutamanya apabila wujudnya ambang seperti buih dan kawasan yang cerah yang mungkin dicam sebagai ikan. Selain itu, sistem pengecaman dan pengenalpastian yang sedia ada masih boleh ditingkatkan untuk mendapatkan keputusan yang lebih baik dan tepat. Dalam usaha untuk mencapai kadar pengecaman dan pengenalpastian yang lebih baik, satu skim yang ditingkatkan yang terdiri daripada gabungan beberapa kaedah dibangunkan. Pendekatan pertama terdiri daripada satu kaedah pengesanan objek untuk mencari lokasi ikan. Sebuah sistem menggunakan dwikamera juga dicadangkan untuk memperoleh kadar pengesanan yang lebih baik. Ini termasuk pertimbangan untuk mengesan pelbagai ikan dalam sebuah tangki ikan secara berautomatik. Kadar pengesanan dan pengenalan mungkin lambat kerana proses berjalan secara berterusan terutamanya dalam persekitaran masa nyata. Satu kaedah pengesanan ikan yang lebih tepat dan kaedah yang sistematik untuk mengenal pasti dan mengesan pola renang ikan dicadangkan. Dalam kajian ini, algoritma penyaring partikel dipertingkat dan seterusnya digabungkan dengan algoritma pengesanan gerakan untuk tujuan pengesanan ikan. Pendekatan kedua termasuk reka bentuk dan kaedah pembangunan yang dipertingkatkan untuk memangkas dan mensegmentasikan imej secara dinamik dalam persekitaran masa nyata. Kaedah ini dicadangkan untuk mengekstrak setiap imej ikan dari setiap turutan bingkai video untuk mengurangkan kecenderungan algoritma berkenaan mengesan latar belakang sebagai sebuah objek. Pendekatan ketiga adalah penggunaan kaedah pencirian objek yang disesuaikan menggunakan pemerihal ciri warna untuk mewakili ikan dalam bentuk komputan bagi proses seterusnya.

Dalam kajian ini, sebuah kaedah pencirian objek, GCFD (*Generalized Colour Fourier Descriptor*) disesuaikan kepada persekitaran berkenaan untuk pengenalpastian ikan yang lebih tepat. Kaedah pemadanan ciri berdasarkan pemadanan jarak digunakan untuk memadankan vektor ciri imej yang disegmentasikan untuk tujuan mengelaskan ikan yang spesifik dalam video yang dirakam. Di samping itu, sebuah sistem prototaip yang memodelkan pola renang ikan yang terdiri daripada semua kaedah yang dicadangkan dibangunkan untuk menilai kaedah-kaedah yang dicadangkan dalam kajian ini. Berdasarkan keputusan yang diperoleh, kaedah-kaedah yang dicadangkan menunjukkan keputusan yang lebih baik yang seterusnya menghasilkan sistem pengecaman dan pengenalpastian ikan dalam masa nyata yang lebih baik. Kaedah pengesanan objek yang dicadangkan menunjukkan peningkatan berbanding dengan algoritma penyaring partikel asal. Proses pemangkasan dan segmentasi dinamik dalam masa nyata mencatatkan peningkatan purata sebanyak 84.71%. Kaedah pencirian objek yang telah diadaptasikan untuk pengecaman dan pengenalpastian ikan dalam masa nyata menunjukkan peningkatan berbanding dengan pemerihal-pemerihal ciri sedia ada. Secara keseluruhannya, output utama daripada kajian ini boleh digunakan oleh penternak ikan untuk mengesan dan memantau ikan di dalam air secara automatik berbanding dengan cara konvensional.

AUTOMATED FISH DETECTION AND IDENTIFICATION

ABSTRACT

Recognition and identification of fish using computational methods have increasingly become a popular research endeavour among researchers. The methods are important as the information displayed by the fish such as trajectory patterns, location and colour could determine whether the fish are healthy or under stress. Current methods are not accurate especially when there exist thresholds such as bubbles and some lighted spots which might be identified as fish. Besides, the recognition and identification rate of the existing systems can still be improved to obtain better and more accurate results. In order to achieve a better recognition and identification rate, an improved scheme consisting of a combination of several methods is constructed. First of all, the first approach is to propose an object tracking method for the purpose of locating the position of fish for real-time videos. This includes the consideration of tracking multiple fish in a single tank in an automated way. The detection and identification rate may be slow due to the on-going tracking process especially in a real-time environment. Real-time refers to the ability to perform the detection, recognition and identification process providing an output without significant delay. A more accurate fish tracking method is proposed as well as a systematic method to identify and detect fish swimming patterns. In this research, the particle filter algorithm is enhanced and further combined with the motion detection algorithm for fish tracking. A dual camera system is also proposed to obtain better detection rate. The second approach includes the design and development of an enhanced method for dynamically cropping and segmenting images in real-time environment. This method is proposed to extract each image of the fish from every successive video frame to reduce the tendency of detecting the background as an object. The third approach includes an adapted object characterisation method which utilises colour feature descriptors to represent the fish in a computational form for further processing. In this study, an object characterisation method, GCFD (Generalized Colour Fourier Descriptor) is adapted to suit the environment for more accurate identification of the fish. A feature matching method based on distance matching is used to match the feature vectors of the segmented images for classifying the specific fish in the recorded video. In addition, a real-time prototype system which models the fish swimming pattern incorporating all the proposed methods is developed to evaluate the methods proposed in this study. Based on the results, the proposed methods show improvements which result in a better real-time fish recognition and identification system. The proposed object tracking method shows improvement over the original particle filter method. Based on the average percentage in terms of the accuracy for the dynamic cropping and segmentation method in real time, an acceptable value of 84.71% was recorded. The object characterisation method which is adapted for fish recognition and identification in real time shows an improvement over existing colour feature descriptors. As a whole, the main output of this research could be used by aquaculturist to track and monitor fish in the water computationally in real-time instead of using the conventional way.

CHAPTER 1 - INTRODUCTION

Object tracking using computational methods has become a popular research endeavour among researchers. However, most research involves tracking of human while not much work has been carried out on tracking of animals such as fish. Different concepts have been introduced in the tracking of fish such as installing water sensors and video cameras to identify movement speed, colours, shapes and swimming patterns of the fish. Thus, tracking of the fish poses a great challenge. The main problem is that the fishes swim in various directions and angles with variable speed. So, it is difficult to track and identify a specific fish in a school of fish in most cases. Besides, the fish may also appear to be overlapping one another.

Object characterisation which is an important component of object recognition has become a popular research area over the last two decades. Industries ranging from food and electronics often utilise object recognition techniques and apply them in machines for packaging and categorising items. Based on some observations, it is found that it is important to study specific characteristics of an object for better recognition. For example, fish may come from different types of species. Some fish can be categorised as active swimmers and some as inactive swimmers. If the characteristics of the fish can be detected, the species may be identified and they provide additional information to the characterisation process.

The characteristics of an object mainly refer to the features of the fish which are often calculated in greyscale. However, this research work considers colour video frames instead of greyscale images as colour images provide additional information for the characterisation process. Indeed, colour images provide vectorial data and are more difficult to process than the scalar values of the greyscale images. The benefit of using colour images rather than greyscale/binary images is that a unique set of features for each fish can be provided. Colours also provide valuable information for object description and matching.

Therefore, this research consists of exploring a mechanism to track and characterise fish which are recorded in real time. Obtaining a better identification and recognition rate is the main intention of this research.

1.1 Object Tracking

Object tracking algorithm focuses on identifying the location of an object in order to further obtain other information such as trajectories and general behaviour. Several popular object tracking algorithms include motion detection algorithm and particle filter algorithm (Jacek, 2006). These two algorithms are commonly used in surveillance systems.

Motion detection enables the detection of moving objects in two dimensional or three dimensional environments (Kirillov, 2007). It is often used in surveillance and monitoring systems for tracking moving objects in uncontrolled scenarios such as places with no proper lighting and in real-time situations (Sawhney et al., 2000; Hyenkyun et al., 2010; Weiming et al., 2004). Many different concepts have been applied in motion detection (Ma, Chi, & Zhang, 2004; Jing et al., 2005). Motion detection can be applied to video capture frame by frame in order to identify moving objects. A movement in the video indicates that there are activities and the level of activeness is detected by using a sensitivity parameter. The sensitivity level is generated by comparing successive frames in the video. A high sensitivity level indicates the existence of a very aggressive movement in the video while a low sensitivity level shows that there is either no movement or very little movement.

Particle filter algorithm uses a weighted distribution approach which is propagated through time using a set of equations known as the Bayesian filtering equations. The trajectory of the tracked objects is determined by taking the particle with the highest weight or weighted mean of the particle set at each time-step (Islam et al., 2009). This algorithm consists of three steps in the implementation process which are the initialisation step (prediction), sampling step (update) and selection step (resample) (Mariño, et al., 2011). Particles are initially scattered across a region of interest in the image. For every time-frame, the object which moves will be identified based on a motion or a colour model. The particles around the moving object will be given higher weight. As the object continues to move, the particles will accumulate around the object. This enables tracking of the trajectory patterns and location of the object based on the highest weight at each time-step.

Condensation algorithm is an extended version of the particle filter algorithm (Koller-Meier & Ade, 2011). It is able to track objects which are represented by parameterised spline curves in substantial clutter at video frame rate. The unique feature of this algorithm is that it does not compute every pixel in an image frame. Instead, the pixels to be processed are chosen randomly. Besides, condensation algorithm is able to run in real time on modest hardware. This algorithm has been applied in tracking of multiple objects and often used in traffic scenarios because vehicles move regularly on the road (Meier & Ade, 1999; Koller-Meier & Ade, 2011). Lui et al. (2008) used an adaptive approach to define condensation algorithm in modelling a face tracking problem.

Motion detection can be used in tracking of movements in a real-time environment. Therefore, one of the methods to track fish activity in the water is a method that uses motion detection algorithm. The level of information which can be obtained is just the swimming activity which is based on movements. The problem with motion detection is that a level of activeness has to be specified and if the objects move slowly, the motion detection algorithm may not work. Many researchers used the particle filter algorithm in the tracking domain but most of the applications only cater for a single object. Handling overlapping objects is also a difficult task because two or more overlapping objects maybe mistracked after a certain period.

Therefore, there is a need to further improve the object tracking algorithms in order to provide a better way of tracking fish in real-time video.

1.2 Object Characterisation

Object characterisation mainly refers to the recognition of items or animals either in a static environment or in a real-time environment. In this context, the terms object characterisation and object recognition are used interchangeably. Object characterisation normally comes after the object tracking process as specific objects are required to be tracked before information can be extracted.

Nowadays, object characterisation involving human beings, animals or plants has become an interesting area of research (Saravanakumar et al., 2010; Nan et al., 2011). Various methods are used regardless of hardware and software to characterise and classify objects. Object characterisation method is widely used in the domain of information retrieval, object classification and object tracking (Wang et al., 2005). A feature descriptor represents the points which are detected in an image as feature vectors. The values in the feature vectors differ from one to another depending on the type of feature descriptor used. Each feature vector keeps the unique information which can be used to match an object or for information retrieval based on images. Therefore, there is a need to improve the current object characterisation methods to provide a better recognition method for specific objects, mainly fish.

1.3 Fish Recognition and Identification

The purpose of recognising fish is to identify the location of every specific fish from a school of fish. Different types of fish display different characteristics such as swimming patterns, features and colours. Hsiao et al. (2013) used a sparse representation to recognise and identify the fish utilising the maximum probability of the partial ranking method.

Wright et al. (2010) used a sparse signal representation to identify the signals which contain semantic information of an image. The representation of sparse signals can be in the form of a wavelet or Fourier transform. The ability to transform an image into a different form of representation (dimension signals) such as the sparse representation enables further image processing techniques to be applied.

Some researchers used the method of eigenfaces and fisherfaces to extract feature data. They are widely used in face recognition (Belhumeur et al., 1997; Yang, 2002; Turk & Pentland, 1991). The reported results are good in terms of the recognition rate (81.8%) and the identification rate (96%). The authors use a set of databases consisting of different species of fish. The parameters set by the authors are the dimension of the feature space and the partial ranking value.

Therefore, this research will also focus on improving fish detection and identification method for the aquaculture community.

1.4 Motivation

Nowadays, object tracking, and particularly tracking of human beings, animals or plants has become an interesting area of research. Various methods are used regardless of hardware and software to improve the effectiveness of the tracking algorithm. Previous studies have proven that a great amount of work has been done on object tracking. Human beings have a limitation as far as tracking is concerned compared to the computational methods used in tracking. In this research, the object referred to is mainly fish. Fishes in this world can be divided into species, which appears to have different sizes and colours. Fishes are commonly taken as food. So, in order to maintain the quality of the fish, constant monitoring is required, but it is not possible to monitor the fish all the time. Moreover, the required knowledge on fish behaviours is vast in order to classify whether the fish are in healthy or unhealthy.

Therefore, the main motivation of this work is to design and develop an enhanced fish tracking method based on the identified fish swimming patterns. Besides, a better identification and recognition of fish is proposed in order to further classify each fish swimming in the water.

1.5 Problem Statement

Existing fish tracking systems can still be improved mainly in the domain of object detection and identification (Jia-Hong et al., 2010). Not many fish tracking systems are available in the market due to the low accuracy in the tracking process of the fish. Existing tracking systems depend solely on hardware and are geared more towards simulation than for real-time environment (Shao & Xie, 2012). Besides, to our knowledge, there is no study where dual cameras are used in fish tracking. Most studies are concerned with the use of a single camera to monitor one whole fish tank (Chew et al., 2009; Jia-Hong et al., 2010). Therefore, there is a need for an enhanced object tracking algorithm to provide a better detection rate. Using a dual camera system provides additional information from the side view and top view so that the accuracy of the fish tracking system can be improved further especially for overlapping fish. By having dual camera system, there is a possibility to track fish using the additional information generated from the top view to further improve the tracking accuracy.

Moreover, the image sequence may contain noise (due to external factors) and may appear cloudy. The existing fish tracking methods are not accurate enough such as in the case where some external factors (such as bubbles and some lighted spots) might be identified as fish (Gracias et al., 2008; Feifei et al., 2011). Noisy and cloudy images also contribute to the increasing difficulty of the detection process. As a result, image cleaning using noise reduction software is important as a pre-processing tool (Inamori et al., 1993). Therefore, there is a need to pre-process images to remove, or at least reduce external factors while retaining the required object in the image for further processing by implementing an image segmentation process for extracting the required object.

The recognition rate and identification rate can be further improved especially in real-time environment. The problem of fish recognition and identification is difficult due to the continual changes in movement, swimming angle of the fish and the condition of the fish which indicates whether the fish is healthy or unhealthy (Wang et al., 2010). Besides, the variety of fish together with the coral reefs and water plants present a challenge in the recognition process (Chen, 2003). A lot of studies on fish recognition and identification system are applied only on static images due to heavy computation (Rodrigues et al., 2010). Therefore, there is a need to adapt the existing object characterisation methods to suit specific scenario which are the characterisation of fish in real-time videos in order to obtain better recognition rate and identification rate.

Therefore, this study proposes the use of object tracking algorithms and colour object recognition technique to detect the change in speed or vector movement (trajectory) of the fish in the viewing field. The objective of this method is to enable the fish to be tracked down digitally. The required objects to be recognised in this research are the fishes. Therefore, the research questions are:

- How to enhance the existing object tracking method for detecting the location and trajectory patterns of the fish in real-time videos?
- How to improve the quality of image cropping of real-time videos for better detection and recognition of the fish?
- What adaptations are required to the existing classification and characterisation methods for fish characterisation and a better level of detection and recognition rate of fish in real time?

1.6 Objectives

The aim of the research is to provide a better automated detection method for identifying fish swimming patterns for real-time videos and characterise the fish based on an improved characterisation method. Therefore, in order to achieve this goal, the objectives of this research are as follows:

- To enhance the object tracking method for a more accurate detection of the location and trajectory patterns of the fish (fish swimming patterns).

- To design and develop an enhanced method for dynamically cropping and segmenting images in real-time which can be applied to video frames.

- To adapt the object characterisation method for a better detection and recognition rate of the fish.

- To develop a complete real time fish identification and recognition system based on the above proposed methods.

1.7 Research Contributions

The contributions of this research are as follows:

- An improved object tracking algorithm based on particle filter algorithm for tracking fish.
- A more accurate fish tracking method to detect fish swimming patterns based on the location and trajectories of the fish. The enhanced fish tracking method considers several methods which include:
    - An improved object tracking algorithm that combines the motion detection algorithm and the particle filter algorithm (or an enhanced particle filter algorithm).
    - An improved way of tracking overlapping fish by utilising dual cameras positioned on the side and top of a fish tank.
- An enhanced dynamic cropping and segmentation method for real-time videos.
- A more accurate fish recognition and identification method for matching each specific fish based on an adapted colour feature descriptor which considers a transformation of an image in a different representation.
- An improved scheme for better detection and recognition rate for tracking of fish in real time using the above mentioned proposed methods.

1.8 Scope and Limitation

The scope of this research is restricted in several ways. For instance, this research is carried out on fish in a cultured fish tank where water quality such as the salinity, oxygen level and temperature of the water can be manually adjusted. The size of the fish tank is six feet in width, two feet in depth and three feet in height. The size of the fish tank only affects the density of the fish in the fish tank. The number of fish which occupies the fish tank is more important in measuring the swimming patterns of the fish. In order to have a wider range of datasets which contains different densities of fish, additional datasets from FishCLEF (ImageCLEF - Image Retrieval in CLEF, 2014), mainly underwater videos, are also used for testing and evaluation.

Several attributes corresponding to the fish such as colour, trajectories and location are considered. This study focuses on developing a fish monitoring system mainly for the aquaculturist to track and identify whether the fish in a fish tank are healthy or not. The experiment is conducted using two video cameras mounted in a cultured fish tank containing several koi fishes. Koi fishes are used due to the active swimming behaviour in the water. Nonetheless, the method can also be used for slow moving fish. The videos are captured from two points of view (top and side views) which can produce a single camera video by removing the top view. For FishCLEF datasets, only videos using a single video camera are available. However, this dataset will still be used for the evaluation since it is widely and publicly used by other researchers in this area.

The research is carried out to perform in a real-time environment. Real-time refers to the ability to perform the detection, recognition and identification process, providing an output without significant delay. The processes will be automated where no human intervention is required. Therefore, a high-speed computer system which is able to execute all the processes is used.

The focus of this research consists of three main methods which are object tracking, dynamic cropping and segmentation, and object characterisation. The final result from the characterisation process could be used by the biologist to classify whether the fish is healthy or unhealthy. This research will be done up to the point where the location of the fish and the characteristics of the fish are detected. The work of the biologist in analysing the fish behaviour will not be considered in this research.

1.9 Benefits of the Research

This research benefits mainly the aquaculture industry. The benefits of the research are as follows:

- This research will give a big boost to the fish culture industry in Malaysia especially for exotic fish such as arowana, koi and flower horn. Besides, food industries which cultivate fish for human consumption will also benefit from this research.
- Aquaculturists and fish breeders could save money by reducing the use of costly hardware for monitoring the fish ponds. Human errors will also be reduced and the fish can be monitored all the time.
- This work could be extended to the prawn farming industry. Prawn farming requires proper hatcheries with good water quality, climate, labour and so on. By adjusting the required parameters, this research can cater to a larger variety of industries locally in Malaysia and globally.

1.10 Research Approach

An overview of the process involved in this research is shown in Figure 1.1. This process consists of several stages where each stage represents a task in fulfilling the objectives of this research.

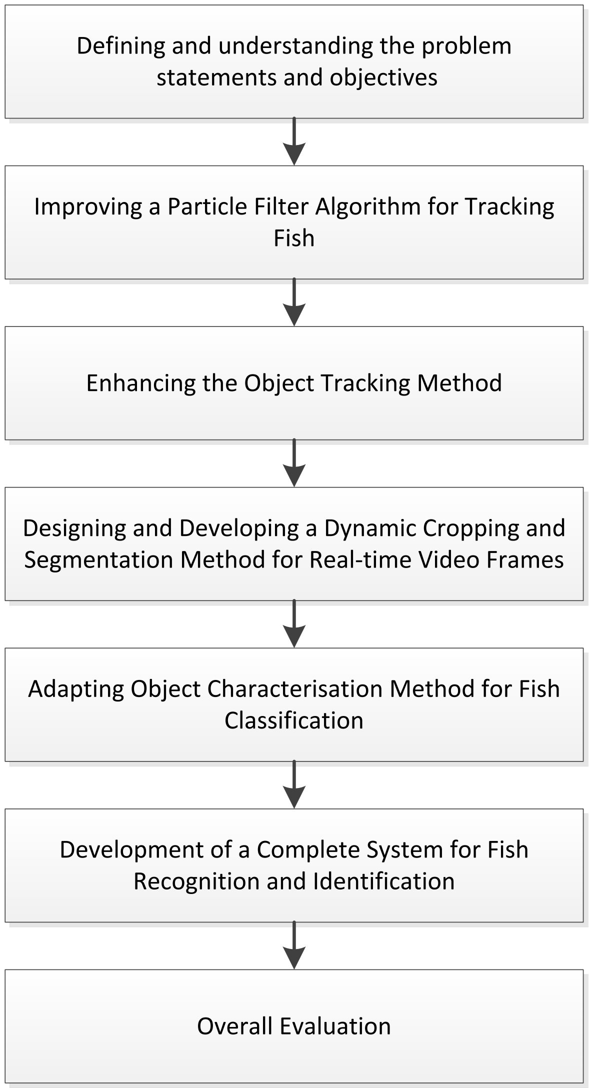

Figure 1.1: Overview of the Research Approach

This research begins by defining and understanding the problem statement and objectives of the research. This stage which involves consultation with the biologists is carried out to define specific objectives of the proposed research. The case study of the research is conducted in a lab located at the School of Computer Sciences, Universiti Sains Malaysia. The stage consists of developing a research plan in which necessary studies and timeline are designed, and the required methods are identified and studied.

The next steps in this research consist of producing an enhancement in object tracking method, developing an automated dynamic cropping and segmentation, and adapting fish characterisation method for fish classifying in real-time videos. An improved particle filter algorithm for tracking fish is proposed in this research. This algorithm is further enhanced to provide a more accurate object tracking method. The result from the improved particle filter algorithm will be compared with the existing particle filter algorithm to evaluate the results of both algorithms using precision and recall method. For the object tracking method, the precision and recall evaluation method is used. After having an enhanced object tracking method, an enhanced image dynamic cropping and segmentation method for video frames in a real-time environment is implemented. The output from the image dynamic cropping and segmentation method will be a set of images containing the fish that provides an input for the object characterisation method. Dynamic cropping and segmentation method will be evaluated by comparing the manually cropped images with the dynamically cropped images. The total number of generated images based on the cropping process will be recorded and the accuracy of the segmentation process is calculated. In order to improve the identification and recognition of fish, an existing object characterisation method is adapted. A complete system for fish recognition and identification is developed for testing and evaluation of the enhanced object tracking method and the enhanced methods for object characterisation.

The final stage of this research is the overall evaluation which in effect is the evaluation of the object characterisation method. The evaluation is carried out between different keypoint detectors and feature descriptors where a set of descriptors are generated from different methods of feature descriptors (Sande, 2010). These descriptors will then be analysed further using *k*-means and histogram distance. Several feature descriptors have been evaluated to compare the accuracy with the adapted feature descriptor proposed in this research.

Three datasets (two own datasets and an existing dataset) will be used in this research. The first dataset consists of using two video cameras recorded in a cultured fish tank which captures the side view and the top view of the fish tank. The reason for producing the first dataset is because there is no existing videos recorded using dual video cameras for fish recognition and identification. The second dataset is obtained from the first dataset by removing the top view of the recorded video. The third dataset comes from a series of video captured in an underwater environment called FishCLEF. In this video, only a single camera is used. The videos are updated every year to provide a benchmark for other researchers to measure the effectiveness of their algorithms or methods.

1.11 Thesis Organisation

In the next chapter (Chapter 2), a brief literature review concerning object tracking, image cropping and segmentation, and object characterisation is described. The chapter also includes a review on existing methods of fish recognition and identification. A summary of all the existing methods is also presented.

Detailed explanation of the enhancement process involved in the object tracking method is described in detail in Chapter 3. This includes the improved particle filter algorithm and a combination between the original particle filter (or improved particle filter) with the motion detection algorithm. Chapter 3 also describes the construction of the datasets using single and dual camera. The evaluation of the enhanced object tracking method is also explained in this chapter.

The algorithm for the improved dynamic cropping and segmentation process is explained at the beginning of Chapter 4. The dynamic cropping process is improved to enable the method to be executed in real-time environment. At the same time, an image segmentation method is incorporated to remove the background from the object which is explained in the same chapter. Then, the adaptation process in the object characterisation method will be further explained which provides an overview of the enhanced methods for object characterisation. The evaluation of the enhanced object characterisation is also explained in the same chapter.

The development of the complete system of fish identification and recognition is detailed in Chapter 5 which includes the implementation process and application of the system. This chapter also explain in detail the training process (learning phase) and the testing phase in detail for fish identification and recognition.

The results and discussion are presented in Chapter 6 which includes the evaluation results of the improved object tracking method, improved dynamic cropping method and the improved object characterisation method. These methods are evaluated using two own datasets (dual and single camera) as well as existing dataset. Chapter 7 concludes the research.

CHAPTER 2 - LITERATURE REVIEW

In this chapter, several existing methods used in object tracking, object recognition and object identification are reviewed. Then, a more specific review on fish recognition and identification methods are carried out. As a whole, this chapter will further explain the existing methods which are used in recognition and identification of specific objects. The objects are mainly referred to as fish since the objectives of this research are to track, recognise and identify the fish swimming in a fish tank.

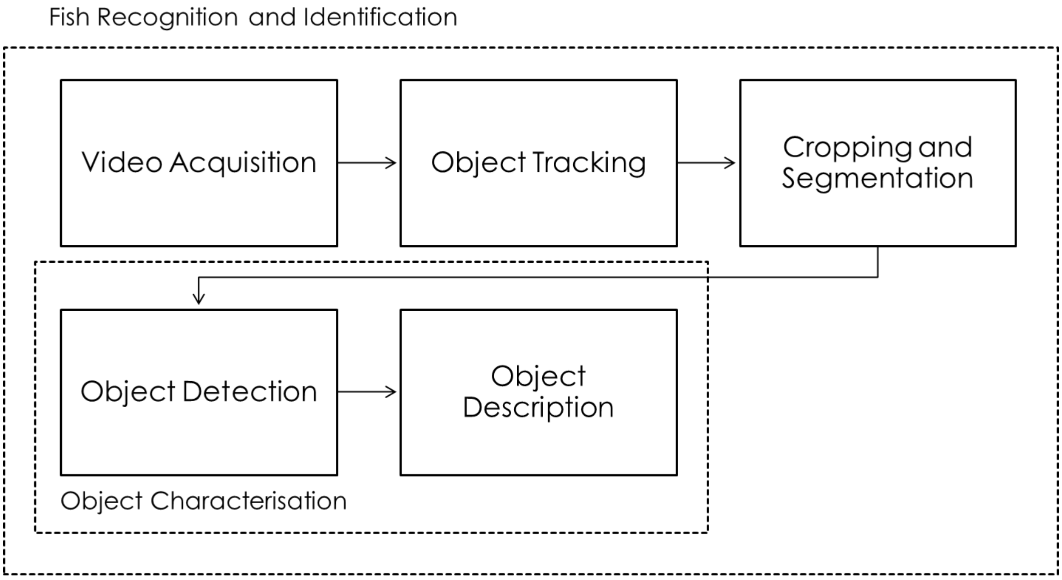

Figure 2.1: Global Scheme of Research

Figure 2.1 shows a global scheme of the research. The scheme consists of five different blocks which are the video acquisition, object tracking, cropping and segmentation, object detection and object description. The object detection and object description blocks are grouped as object characterisation. The video acquisition is not going to be reviewed in this chapter. The object tracking will be reviewed to resolve the issue where existing fish tracking systems can still be improved mainly in the domain of object detection and identification. Moreover, the image sequence may contain noise and may appear cloudy that should be resolved using the cropping and segmentation method. The object characterisation is then reviewed to resolve the issue where the recognition rate and identification rate can be further improved especially in real-time environment. Combining all the blocks provides a complete fish recognition and identification system. These blocks will be reviewed in this chapter to study the current issues in these methods.

Section 2.1 consists of a set of literature reviews on object tracking. Object tracking consists of various methods which specifically tracks an object or for surveillance. This section classifies the object tracking methods into several categories which are motion detection algorithm, particle filter algorithm, real-time mean-shift tracker, blob detection and template based object detection.

Section 2.2 describes several existing methods related to image cropping and segmentation. This section briefly reviews the existing work carried out by researchers on image cropping methods and image segmentation methods.

As this research relates to the recognition and identification of fish, a brief review on existing work in this area is carried out and explained in Section 2.4. This section reviews the work on the aspect of using fish as one of the components in the research. Therefore, the study includes a brief literature review on the colour changes on fishes and the study on fish swimming patterns.

2.1 Object Tracking

Object tracking has become a popular and challenging research domain due to the various changes displayed by the object such as movement, colour, speed and shape. Different objects exhibit different changes. The objects refer to non-static entities such as human beings or animals. The challenge that arises in object tracking is that several objects mainly with the same characteristics such as colours and shapes appear in the same video frame. There are several tracking methods such as gradient-based image flow, feature-based image flow and mean shift tracking. Each of these methods has their own advantages and disadvantages depending on the input data and scenario of an application.

Meng-Che et al. (2013) proposed a tracking method by utilising the Viterbi data association method. However, the proposed method only works well with proper segmentation of the object. There is no consideration in terms of background objects which may degrade the tracking rate.

Several popular object tracking algorithms such as motion detection algorithm and particle filter algorithm are studied in order to identify the characteristics and techniques used to track objects. These two algorithms are commonly used in surveillance systems. Object tracking algorithm focuses on identifying the location of the object in order to further obtain other information such as trajectories and directions. These two algorithms are reviewed in the following section.

Background subtraction method has been used as a method for tracking objects. It is a method applied in pre-processing of video frames in order to remove background while retaining the objects in the video frame. This method provides a good way of tracking objects but some tracking errors still occur when objects overlap or there are other side effects such as lighting and noise appearing in the video frame (Rahman et al., 2009).

Zhong et al. (2000) introduced a new method for object tracking using prototype-based deformable template models. In order to track the objects in successive frames, the authors modelled the shape of the object. Three categories of image information are used in the method namely considering the edge and gradient, region consistency and interframe motion. However, this method is relative to the shape of the object and is unable to perform in situations where images are deformed. Besides, multiple object tracking is restrained in this method.

Tracking multiple objects can be done by identifying the objects of each successive frame and connecting the detected objects of the frames. Consequently, if an object is detected in the first frame but not in the following frame, the trajectory of the object may not be accurate. Berclaz et al. (2011) mentioned that multiple target problem is usually solved using sampling or greedy search based on variants of Dynamic Programming. However, it may not be practical as it can be excessive in processing. Therefore, the authors introduced k-shortest paths algorithm in object tracking. In the algorithm, it can track objects at a moderate speed, but if the objects are moving at high speed, the algorithm may not perform well in object tracking.

Object tracking can be regarded as a difficult task which could be affected by noise or illumination in a video. Gevers (2004) mentioned that object tracking in video is important for video-based encoding, surveillance and retrieval. Therefore, a method based on colour-based deformable model which distinguishes noise and illumation is proposed. Good results from the object tracking process were obtained based on the experimental data as shown in Figure 2.2. However, the ground truth data used by the author shows a high level of colour difference (distinct colours) between the objects and background which makes the tracking process attain higher accuracy. Besides, the colour of the object has to be set for the tracking process.

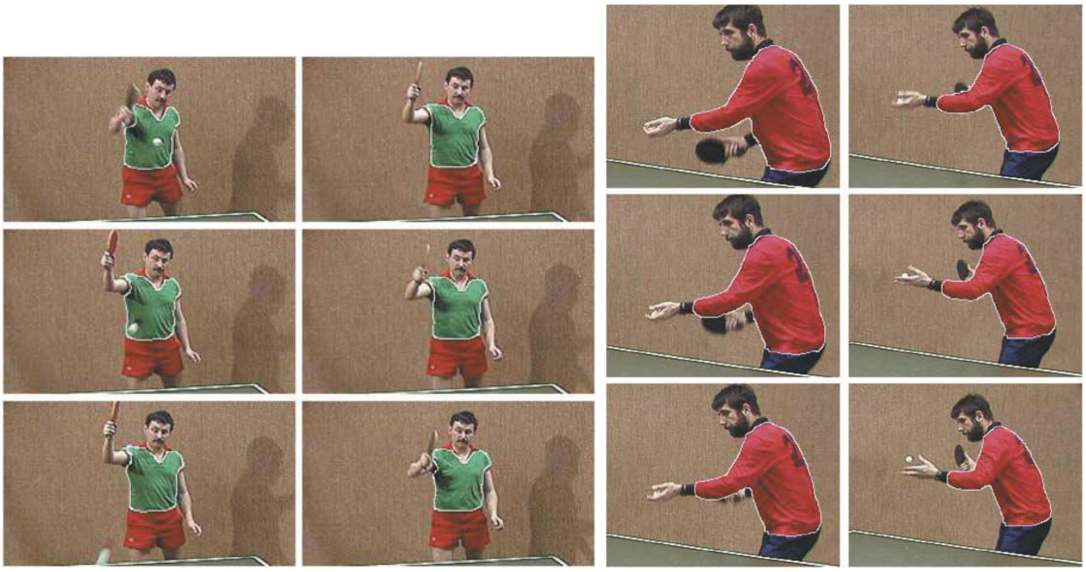

Figure 2.2: Tracking of Object Based on Colour-based Deformable Model (Gevers, 2004)

Koo et al. (2013) proposed a novel method for multiple object tracking using Hierarchical Spatiotemporal Data Association (HSTA). The method allows tracking of multiple objects while allowing the correction of wrongly detected objects even with the presence of other objects. This method performs well in tracking and associating between different types of object as shown in Figure 2.3. However, in the test case, there isn't any diverse colour background which may affect the outcomes of the tracking process. This approach requires a pre-processing algorithm to identify the objects before the tracking of the movements is carried out.

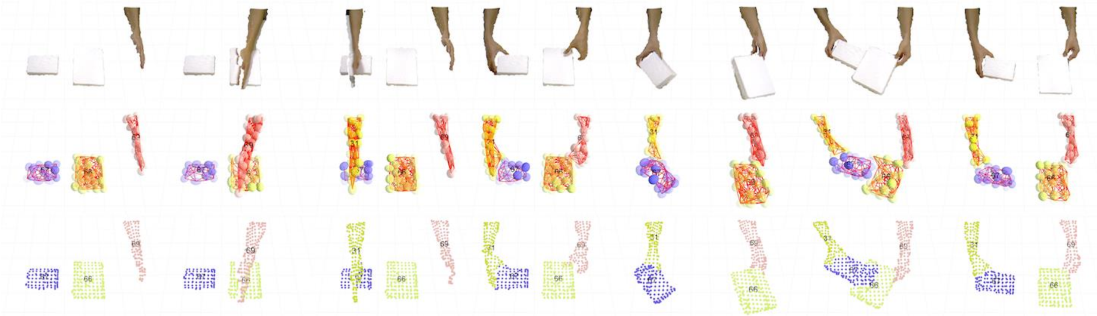

Figure 2.3: Snapshots of Tracking Multiple Objects in a Sequence of Movements (Koo et al., 2013)

2.1.1 Motion Detection Algorithm

Motion detection is applied to the captured video frame by frame in order to identify moving objects. A movement in the video indicates that there are activities and the level of activeness is detected by using a sensitivity parameter. The sensitivity level is generated by comparing successive frames in the video. A high sensitivity level indicates the existence of a very aggressive movement in the video while a low sensitivity level shows that there is either no movement or very little movement (Denzler & Paulus, 1994).

However, based on a brief literature review of the available studies, a lot of motion detection algorithms are hardware based while processing are done in the hardware itself. Software based motion detection are rare and mainly for the purpose of security. For example, motion detection algorithms are installed in surveillance camera to track if movement exists before enabling the recording mode (Thomas et al., 2011).

Studies on motion detection are mainly used to identify whether movement of a specific object exists in a video. Issues arise when movements of certain objects not under consideration is detected. Existing motion detection algorithm may be used in the form of object tracking but it should be accompanied with some enhancements on the algorithm in order to further improve the accuracy for the purpose of using it in object tracking.

Motion detection is often used in surveillance and monitoring systems for tracking moving objects in uncontrolled scenarios such as places with no proper lighting or real-time situations. In Sawhney et al. (2000), an independent motion detection based on plane-plus-parallax decomposition is proposed. Constant or fast moving objects generate a lot of movements which affect the algorithm's detection of every object. Moreover, continual changes in the background and the generation of noise trigger the algorithm to detect them as objects.

Motion detection algorithms are often embedded in most video surveillance systems which require the sensitivity parameter to be set in order to generate the motion alarm. In the field of tracking, this sensitivity parameter is used to identify the speed of the object. Hyenkyun et al. (2010) proposed a motion detection model based on variational energy which enables a dynamic tuning of the sensitivity level. This algorithm works even in a noisy environment.

Visual surveillance has become a popular research area mainly in computer vision. Weiming et al. (2004) categorised the framework of visual surveillance into several stages. One of the stage is the detection of motion. A motion segmentation is implemented to detect the regions corresponding to the moving objects. The approches used are background subtraction, temporal differencing and optical flow. However, when it comes to a dense environment where multiple fast moving objects are encountered, this method does not provide a reliable detection as the detected shape of the object is inaccurate. The low frames per second affects the detection of motion as well.

Motion detection enables the detection of moving objects in two or three dimensional environments. A motion detection system was developed by Kirillov (2007) which processes each successive image frame of an uploaded video. The motion detection method used in the system enables the detection of objects in a static background. However, overlapping objects generate a lot of false positives when a standalone motion detection algorithm is used.

Many different concepts are used in motion detection for specific purposes. Ma et al. (2004) proposed a new method of using geometric active contour model for motion detection and object tracking. This method is suitable to be applied in a slow changing environment where the object's activity is less occasional.

Jing et al. (2005) proposed a motion detection method by dynamically thresholding the difference between succesive image frames. They uses a histogram method instead of the GMM (Gaussian Mixed Model) method to reduce high computational cost. The problem in this method is that if the object size is not significantly large (small objects), there will be no motion detected. Therefore, the object will be regarded as a non-moving object.

2.1.2 Particle Filter Algorithm

Particle filter algorithm is an algorithm that uses a weighted distribution approach which is propagated through time using a set of equations known as the Bayesian filtering equations. The trajectory of the tracked objects is determined by taking the particle with the highest weight or the weighted mean of the particle set at each time-step (Islam et al., 2009).

In a particle filter algorithm, there are three steps in the implementation process which are the initialization step (prediction), sampling step (update) and selection step (resample) (Mariño et al., 2011). Particles are initially scattered across a region of interest in the image. For every time frame, the object which moves will be identified based on a motion or a colour model. The particles around the moving object will be given higher weight. As the object continues to move, the particles will be accumulated around the object. This enables the tracking on the object's trajectory patterns and location based on the highest weight at each time-step. The issue on particle filter algorithm is mainly on the overlapping of multiple objects which causes inaccuracy or mistracked objects.

The advantages of using particle filter in object tracking are that the algorithm can be easily integrated in a cluttered environment and the implementation is fast and simple. Jacek (2006) proposed the use of particle filter algorithm for object detection. The object in this context is the face. Based on the experimental work, the detection works well on single face recognition but when two or more objects are present, the detection fails. However, after several time frames, the particle filter manages to recover from the detection process.

Okuma et al. (2004) worked on a problem by using the particle filter algorithm for detecting and tracking multiple objects. The main difficulty in tracking multiple objects is in detecting ambiguities and the overlapping objects. The algorithm uses a combination of AdaBoost and the particle filter algorithm. The results from the tracking process were good even when the objects were out of the video frame. However, this method does not work so well when there is a non-uniform background. Therefore, there is a need to further enhance this method for a more robust object tracking.

Kumar and Sivanandam (2012) applied the particle filter algorithm for tracking cars in real time. The authors tracked the trajectory patterns of a moving object using the particle filter algorithm. However, the authors have not applied the technique on multiple objects which are moving at the same time in the same video sequence. This algorithm could be improved by performing some pre-processing on the video frames and considering other parameters such as speed and shape for better detection and characterisation.

Condensation algorithm is an extended version of the particle filter algorithm. The algorithm is able to track objects which are represented by parameterised spline curves, in substantial clutter at video frame rate. The unique thing about condensation algorithm is that it does not take into account all the pixels of an image frame. Instead, the pixels are chosen randomly. Besides, the condensation algorithm is capable of running in real time on modest hardware. This algorithm has been applied to tracking multiple objects and used in traffic scenarios because vehicles move regularly on the road (Meier & Ade, 1999; Koller-Meier & Ade, 2011).

Lui et al. (2008) used an adaptive approach to define condensation algorithm in modelling a face tracking problem. The tracking process in using the condensation algorithm is good but there are some features such as the motion parameter from the motion detection algorithm which could improve the object tracking process.

2.1.3 Mean-Shift based Tracker

Liu and Cheng (2010) proposed an enhanced object tracking method based on mean-shift. The authors also included an occlusion detection method to identify occluding objects as shown in Figure 2.4. The improved method considers the colour histogram model within the tracked area. However, this method is limited to only tracking a single object.

Figure 2.4: Occlusion Detection (Liu & Cheng, 2010)

Yuan-ming et al. (2011) studied the defects of existing mean-shift tracking algorithm and proposed an enhanced method based on evolutive asymmetric kernel to improve the accuracy of the tracking. The authors categorised the method into three steps which are as follows:

1. Introduce an asymmetric kernel function into mean shift algorithm by describing the calculation method of Template Centre.
2. Set the contour evolution algorithm using regional similarity.
3. Perform an update on the asymmetric kernel function.

The results were satisfying as shown in Figure 2.5. However, the method does not work well in real-time environment and is limited to only an object.

Figure 2.5: Racing Car Tracking Results (Yuan-ming et al. (2011))

Gang et al. (2009) mentioned that the mean-shift algorithm is not effective when objects are moving fast because the algorithm does not store the target's motion direction and speed information. The authors proposed a new object tracking algorithm by combining the mean-shift algorithm with motion vector analysis. The movement of the objects are categorised into ten different types (up, down, left, right, left-up, left-down, right-up, right-down, moving away from camera, moving to camera) as shown in Figure 2.6.

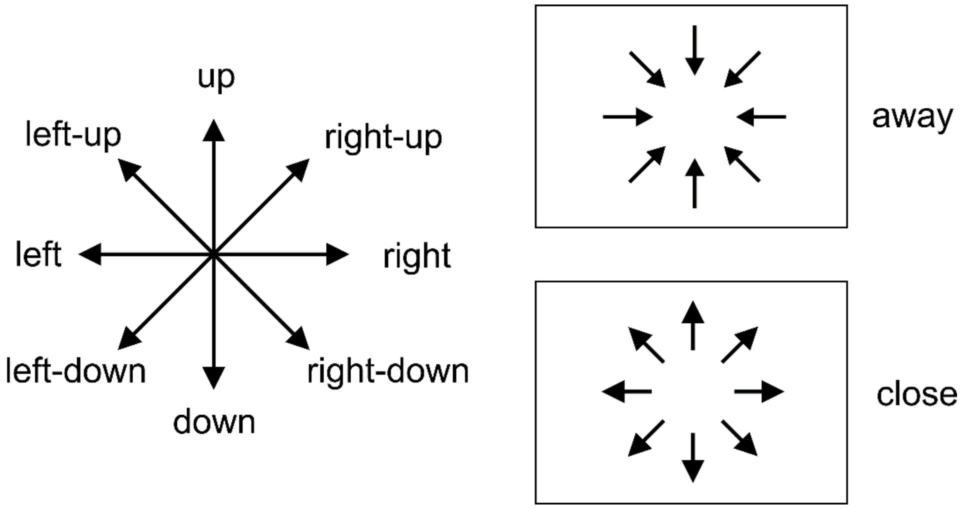

Figure 2.6: Object Moving Direction (Gang et al., 2009)

The tracking process yields a good result and reduces the iterations for faster processing. However, the algorithm does not consider multiple moving objects and overlapping objects.

Tang and Zhang (2011) combined the mean-shift algorithm with the particle filter algorithm to produce an enhanced object tracking algorithm. Mean-shift algorithm is not able to track objects which are moving fast while the particle filter algorithm are able to track objects that are moving fast, but at a slower tracking speed due to the complexity in computation. The novel tracking method is able to track the object in a viewing frame but it is limited to only a single object at a time. The advantage of the method is that it is invariant to rotation and scaling, and has an acceptable tracking speed.

Comaniciu et al. (2000) proposed a real-time mean-shift based tracker used to track objects especially non-rigid ones. Non-rigid object refers to shape which is flexible and does not have a standard shape. However, the proposed method only applies to a single moving object.