A Fuzzy-Mathematical Model to Recover Motion with Monocular Vision

S. Pathirana[1], J. G Samarawickrama[2], A. S. Karunananda[3]
[1]Department of Computational Mathematics, University of Moratuwa, Sri Lanka
[2]Department of Electronic & Telecommunication Engineering, University of Moratuwa, Sri Lanka
[3]Department of Computational Mathematics, University of Moratuwa, Sri Lanka

Abstract– **This paper describes a fuzzy based mathematical model to recover the motion path of an obstacle observed via a single camera. The strategy is analysing a sequence of images captured in regular time intervals, more specifically, studying the variation of the apparent size of an obstacle and the relative position change on reference frame. Those two measures are the only inputs to the fuzzy-mathematical model, the major emphasis of the research, which recovers the motion path as an equation. Necessary feature extraction is being achieved via a real time image processing module which relies on an optical flow technique as the key technique to recognize dynamic objects. It was reported a 91.98% average accuracy from the fuzzy-mathematical model in simulation environments, where the inputs were generated by a simulator program in order to study the precision of the fuzzy-mathematical model as a standalone application. An average accuracy of 59.8% was experienced at the real time application, an artefact to test the postulated concept in real time dynamic environment, which comprised of three major modules: a real-time image processing module, the fuzzy mathematical model and a mobile robot.**

1. INTRODUCTION

Obstacle avoiding is fundamental to the domain of mobile robots. But, coping with dynamic obstacles is absolutely a challenging task since it is extremely difficult to predict the motion behaviour of dynamic objects. This paper describes an attempt to accept the challenge with a fuzzy-mathematical model.

We postulate that a single camera, known as monocular vision in computer vision, is adequate to obtain the perception from the world rather than the widely used technique binocular or stereo-vision. Although stereo vision is appropriate to have the depth information of objects, it is identified as an expensive technique in terms of processing. Nevertheless, it was realized that even with monocular vision, the objective of obtaining the estimated motion path of objects can be achieved, while avoiding the excess burden of aligning and calibrating the two cameras, if stereo vision preferred instead. This advantage was gained due to the innovative concept of using a fuzzy-mathematical model to simulate the motion behaviour, since it requires the only inputs - the apparent size variation of the obstacle in interest and the relative position change of it on the reference frame, both can be measured with only one visual streaming channel.

The fuzzy-mathematical model postulated in this paper produces the equation of the motion path of an object which moves on a straight line, in form of a equation $y = mx + c$, according to a pre-defined 2-D coordinate system. Hence, the primary objective of the research was deriving the relationships among the key parameters of the motion path equation gradient (m), intersection (c) and the variation of apparent size of the obstacle. The milestone was successfully completed when the relationships between those parameters were found as fuzzy membership functions. Then the model designed in postulation stage was refined and fine-tuned with real time experimental data to infer the required outputs: gradient-m and intersection-c of the predicted motion path, once the data set of two variations mentioned above are provided as inputs.

The idea of the FMM (Fuzzy Mathematical Model) was inspired by nature, more specifically, how human judge hazardous dynamic objects seen and react in order to avoid them.

2. RECENT TRENDS IN OBSTACLE AVOIDANCE WITH MONOCULAR VISION

Ashutosh Saxena from Stanford University is a well-known researcher in this domain. A study done by himself and his team [1] was to estimate depth from a single monocular (still) image. They have named it as ' Recovering 3-D depth'. In their research, they also have studied that how human estimate depth from a monocular image. They have initiated with collecting a training set of monocular images (of unstructured outdoor environments which include forests, trees, buildings, etc.). The researchers have used a 3-D distance scanner to collect training data, which comprised a large set of images and their

corresponding ground-truth depth-maps. Their model has used a discriminatively-trained Markov Random Field (MRF). Using this training set already collected, they have trained MRF discriminatively to predict depth.

According to them, their algorithm was frequently able to recover fairly accurate depth-maps.

Although their algorithm appeared to predict the relative depths of objects quite well (i.e. their relative distances to the camera), it has made more errors in absolute depths. They mention that their algorithm appeared to incur the largest errors on images which contain very irregular shapes like trees. Anyway, some of the errors were attributed to errors or limitations of the training set by the researchers.

Our reviews regarding their research are as follows: They have followed a supervised learning approach in order to deal with this problem. The key technique they have used to extract depth information of the image, Markov Random Field (MRF), is a probabilistic distribution analysis approach, more specifically, an undirected graphical model having a set of random variables having a 'Markov property' described by an undirected graph. Since the graph is undirected and subjected to be cyclic, visiting nodes of the graph might be costly due to higher order of connectivity combinations and dependencies. Obviously the model is mathematically complicated and leads to a computationally expensive implementation.

The other fact worth mentioning is that their model has focused on detecting static objects rather than considering dynamic objects. In addition to that, it need to be trained before use, therefore cannot expect functioning properly in previously unseen environments.

Francisco Bonin-Font, Alberto Ortiz and Gabriel Oliver, University of the Balearic Islands, Spain [3] who have done a comprehensive survey on visual navigation for mobile robots point out that the vision-based navigation solutions have mostly been devised for Autonomous Ground Vehicles (AGV), but recently, visual navigation is gaining more and more popularity among researchers developing *Unmanned Aerial Vehicles* (UAV). They also highlight the fact, even a number of solutions for *Autonomous Underwater Vehicles* (AUV) can already be found for many undersea critical applications rely on vision.

They have thoroughly investigated about 40 different research with good-quality results; the interesting observation is 29 out of 40 with satisfactory achievements have been based on single camera, means monocular vision. (That is approximately 72.5%). They have represented the whole three (3) above mentioned domains of applications, ground, UAV and AUV.

They classify the main techniques of vision based object detection as optical-flow-and appearance-based, where Optical-flow-based solutions estimate the motion of objects or features within a sequence of images. They mention that the researchers compute optical flow mostly using (or improving) pioneering techniques from Horn and Lucas and Kanade. In their report, (page 14), they clearly state that vision based systems on optical flow have proved to be especially useful for Unmanned Aerial Vehicles (UAV) because optical flow gives the scene qualitative characteristics that cannot be extracted in detail even from single low quality images.

Yoko Watanabe, states in his Ph.D. thesis [4] that monocular vision based systems have been operating in even more complicated scenarios in UAVs (Unmanned Aviation Vehicles).

He has experimented with UAVs to detect and avoid obstacles while executing a given mission such as preplanned path following or way point tracking and reported satisfactory success in both simulation environments and real-world applications. They have attempted to cope with stochastic behavior of object motion with predictions (page 59). Drawback of the research is the approach of both image processing with Extended Kalman filter (EKF) and estimations based on higher order derivatives in mathematical equations were consuming many computational resources.

The group of researchers lead by Animesh Garg Institute of Technology, India [5] have also done a study on how monocular vision based systems behave in a real world scenario such as vision based obstacle detection and mapping techniques for identifying objects in urban environments for Autonomous ground vehicle navigation. According to them, the applicability of monocular vision has been successfully tested on a test vehicle under variable outdoor lighting conditions and stable results have been obtained. Anyway, their approach was entirely image processing based inference, where recognition of distinct objects done using a primitive method, RGB color separation. But, using ultrasonic sensors for emergency stops in case of proximity to other objects conveys the message that they are not confident of their depth estimation. We interpret it as not a weak point of monocular vision, but inadequate effort of depth estimation and inference. It seems they have experimented

this in controlled environments, rather than busy roads.

Study done by Tobias Low from University of Southern Queensland, Australia and Antoine Manzanera from Ecole Nationale Sup´erieure de Techniques Avanc´ees, Paris, France [6] also implies the fact, ability of monocular vision for object recognition. Although they focus on classification of static objects, an optical flow technique has been employed.

Nicolau Leal Werneck and Anna Helena Reali Costa from Universidade de São Paulo, Brazil [7], Chau Nguyen Viet and Ian Marshall representing University of Kent, Canterbury, United Kingdom [8], from the same research domain, computer vision have also studied use of monocular vision for object recognition, but none of them were interested in coping with moving obstacles rather than classifying and avoiding static objects.

The researchers in a group lead by Jan Hoffmann from Institut f˙ur Informatik, Berlin, Germany have presented a working model [10][11] of use of monocular vision in the RoboCup 2003 obstacle avoidance challenge in the Sony Four Legged League. Their system had enabled the robot to detect even unknown obstacles and reliably avoid them while advancing toward a target. It has used monocular vision data with a limited field of view. Obstacles were detected on a level surface of known color(s). Although system has proved highly successful by winning the obstacle avoidance challenge and was also used in the RoboCup championship games, we cannot accept it as a pure vision based perception due to the reason of using additional sensory devices for 'odometry'. (Odometry is the use of data from moving sensors to estimate change in position over time. Odometry is used by some robots, whether they be legged or wheeled, to estimate (not determine) their position relative to a starting location.)

Above studies generalize the conclusion that, obstacle avoidance task can be achieved via monocular vision, more specifically with image sequence from a video stream. The predictability of motion path of an obstacle with such technique is also confirmed. In addition to that, they promote the technique optical flow to recognize dynamic objects. But none of them have preferred an artificial intelligence approach, hence affected from the drawback, suffering from high computational overhead.

We postulate a fuzzy based modeling is appropriate to model rather than algorithmic or mathematical where the domain complexity and increased computational overhead is unavoidable. The idea is inspired by nature, how human and animals safeguard themselves from moving objects.

Since human are not capable of performing fast mathematical equation solving (i.e. find factors) or even finding the square root of a simple whole number, obviously they do nothing with complicated mathematics such as Markov Random Fields. They never use any odometry sensors, simply use vision as the primary sensory device to perceive the world and do simple inference with less effort.

Attempts for decision making based on analysis of a live video stream with fuzzy logic are rare, but not absent at all.

One of the best example is a study done under the leadership of Rafaelmun˜ Oz-Salinas at Department of Computer Science and Artificial Intelligence, University of Granada. Spain [12] was to recognize doors, distinguish from misguiding objects by shape such as almirahs or large windows for a mobile robot which rely on vision. They mention that the fuzzy logic approach lead to achieve their objectives as expected.

3. SYSTEM OUTLINE

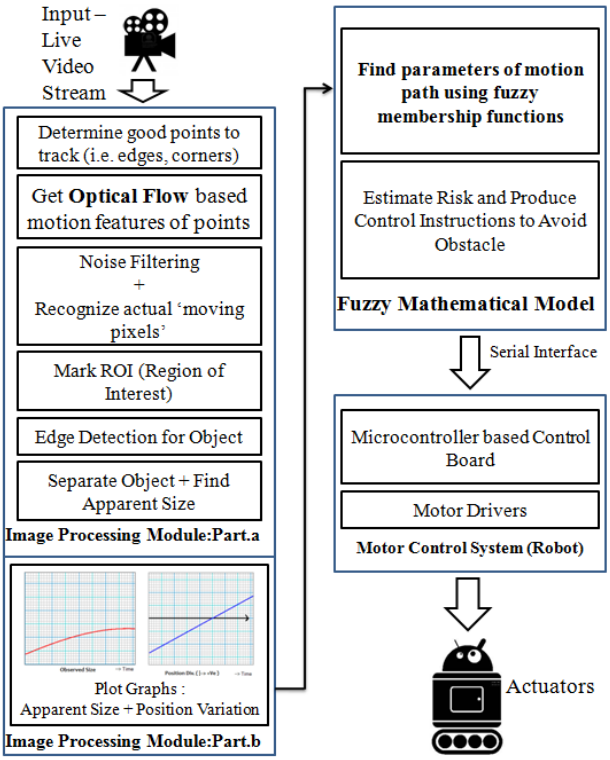

Figure 1: System design overview

The entire artefact is comprised of 3 major modules, namely Image Processing Module, Fuzzy Mathematical Model and The Robot. Task assigned to Image Processing Module is

producing the data sets, the apparent size and position variation, those are the inputs for the Fuzzy Mathematical Model. It again consists of many sub-components of modules, basically for recognizing the target object, after noise filtering. We have improved the existing noise filtering algorithms with density based clustering, which is a statistical approach. Anyway, the key algorithm of this module is an Optical Flow, more specifically Lucas-Kanade (LK) Optical Flow, a Sparse Optical Flow algorithm. The particular optical flow algorithm was selected among many Sparse and Dense Optical Flow algorithms, purely on experimental basis. Dense Optical Flow considers almost all the points to predict the optical flow, which is precise, but was very slow in real time processing. Horn-Schnuck and Gunner Farneback's Optical Flow algorithms are well known. For example, the better one, Gunner Farneback's Optical Flow could process at a rate of 0.25 fps, only 1 frame once in 4 seconds. Therefore, the module was designed with Sparse Optical Flow, which concerns only interesting points of the image, i.e. edges and corners of objects for optical flow predictions. Lucas-Kanade algorithm was selected due to the fast processing capability, which is a must, although it took some time to initialize (track interesting points). We conclude that LK is the most appropriate for real-time processing among existing optical flow algorithms.)

The Fuzzy Mathematical Model, the major emphasis of the research was innovated with the following interesting findings.

One of the most important discoveries was finding the relationship between the gradient (m) of the object's motion path and the apparent size variation. It was observed that the gradient value has some relationship with the skewness (equation (1)) of the curve of the apparent size variation graph, and most probably independent from other factors. This was clearly noticed at fixed points in Table 01.

$$Skewness = \frac{\sum_{i=1}^{N}(x_i - \bar{x})^3}{(N-1)s^3} \quad \text{Where } s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2} \qquad (1)$$

It was noticed that, the skewness value and gradient values are having a relationship in fuzzy nature. But the challenge was to find the actual relationship and below is the strategy followed to determine that.

We constructed a hypothetical fuzzy membership function as Figure.2 (The skewness value was divided in to several ranges and for skewness value between the range *l1* and *l2*) and derived the relationship (3) between the *y* value and some fixed values of gradient *m* as in Table 02.

Table 1: Relationship between the gradient and skewness: Comprehensive

Gradient (m)	Skewness Approx.
-0.1	1.00
-0.2	0.83
-0.3	0.55
-0.5	0.15
-0.6	0
-0.7	-0.14
-0.8	-0.26
-0.9	-0.40
-1.0	-0.52
-1.5	-1.30
-2.0	-3.00

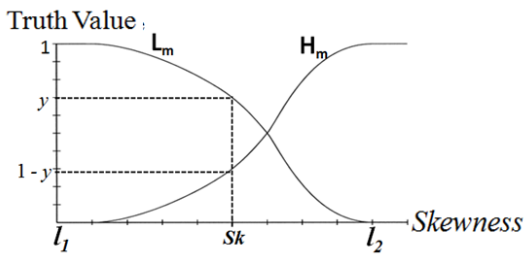

Figure 2: Hypothetical fuzzy membership function for m and skewness

$$m = [y*L_m] + [(1-y)*H_m].(2)$$
$$=> y = (m - H_m) / (L_m - H_m).(3)$$

Table 2: Relationship between the gradient and skewness: Comprehensive

Gradient (*m*)		Skewness (*Sk*)	
H_m(High)	L_m(Low)	*l*1 (lower bound)	*l*2(upper bound)
-0.1	-0.6	1.00	0
-0.6	-1.0	0	-0.52
-1.0	-2.0	-0.52	-3.00

Then, *y* values (*y* axis) (for curve L_m) were found by substituting different *m* values to equation *(3)*, while the corresponding skewness values - *x* axis were experimentally found. A large amount of *x, y* pairs were collected to have a smooth, precise curve as figure 3 therefore the actual fuzzy relationship(s) could be formulated for all skewness value ranges in Table 2 enclosed within *l*1(lower bound) and *l*2 (upper bound). Curve *(1-y)* (for H_m) could be also plotted accordingly (Since y values known). Such an actual fuzzy membership function (When H_m=-1.0, L_m=-2.0, *l*1=-0.52, *l*2=-3.00, refers to last row of Table 2.) is as figure 3.

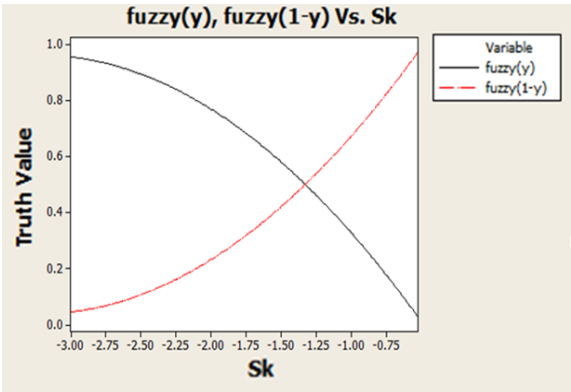

Figure 3: Actual fuzzy membership function of m and skewness

Finding the relationship with the Intersect (c) of the motion path equation and the graph features was another key milestone. This was also successfully achieved with many experiments. Another interesting finding, the relationship between the intersect(c), gradient *(m)* and the initial apparent size(s) was derived.

The experimental findings are graphically illustrated in Figure 4.

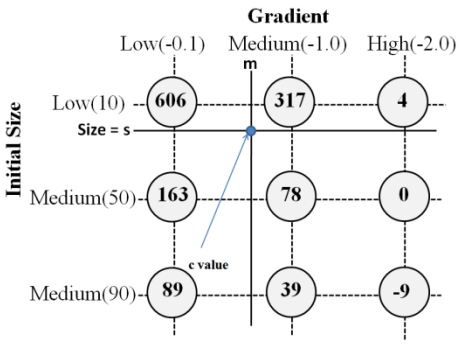

Figure 4: Fuzzy relationship between c, m and initial apparent size(s)

The numerical values within circles are the actual c values at a particular fixed gradient and initial size. For example, the top left most value, 606 is the c value when m=-0.1 and initial apparent size is 10. It was noticed that, the c is varying with m following a linear manner approximately, but c's variation with initial apparent size is non-linear, but maintains a fuzzy relationship.

The variation of c with respect to change of initial apparent size, while keeping m constant, was more scientifically studied. The variation of c, with respect to initial size change, at constant gradients m=-0.1 and m=-1.0 respectively was modeled by Minitab 16 statistical analyzer and was as below:

case 1 (m=-0.1) :

$$c = 849.6 - 32.46s + 0.5116s^2 - 0.002789s^3. \quad (4)$$

case 2 (m=-1.0) :

$$c = 441.0 - 16.65s + 0.2507s^2 - 0.001294s^3. \quad (5)$$

Approximately;

$$c(5) = 2.c(4). \quad (6)$$

This implies that, c value is given by a common function is multiplied with a constant k, i.e. in above case, k = 2.0. This can be mathematically represented as:

$$c = k \,.\, f(s). \quad (7)$$

where k is a constant at when gradient m is maintained uniformly, otherwise k becomes a variable, that is determined by the gradient m. Therefore, in order to find c, it was obvious to find the general function of $f(s)$, then the correlation of k with gradient m. Below is the procedure followed to find the f(s).

Again, hypothetical fuzzy functions were built as Figure 5 and the actual fuzzy relationships were also found on experimental basis in Figure 6.

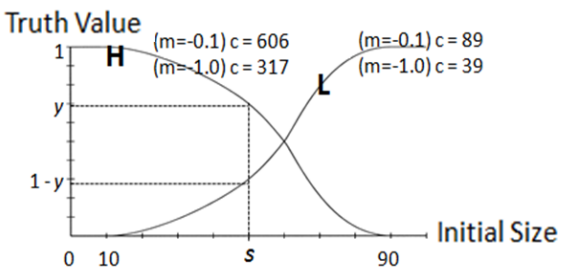

Figure 5: Hypothetical fuzzy membership function of c vs. initial apparent size

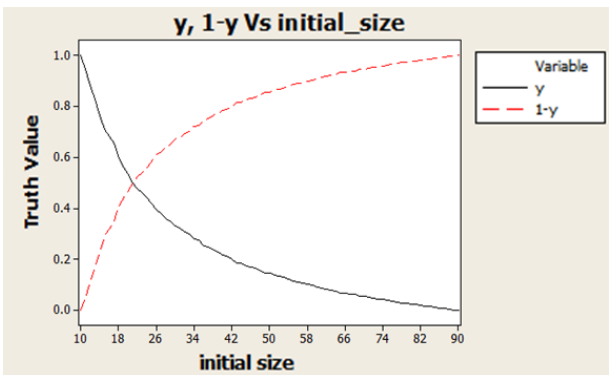

Figure 6: Fuzzy membership function of c with initial size

Then the c value is given by:

$$c = (y * H) + [(1 - y) * L]. \quad (8)$$

Since collected real world data of initial size and *c* value pairs are available, a value matrix of *y* for corresponding initial sizes was found on experimental data, hence the function *y* could be plotted.

As *y* is known from the above fuzzy membership function of *c* with initial size, the two candidate *c* values (for two different fixed gradients) named as c_1 and c_2 is found.

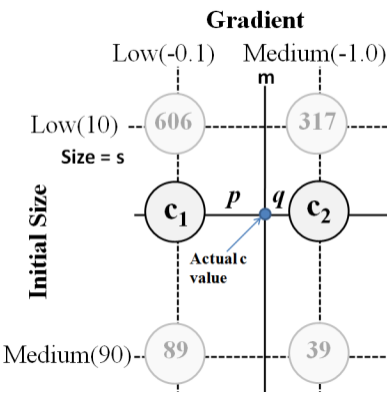

Figure 7: Finding actual c, in between c1 and c2

When the ratio of c_1, c_2 connector line divided by the point "Actual *c* value", (*p*:*q*) is known (assuming relationship between *c* and *m* is linear), final *c* value can be estimated. The final value for intersection *c* is given by the fuzzy membership function specified in Figure 8 as its *y* value.

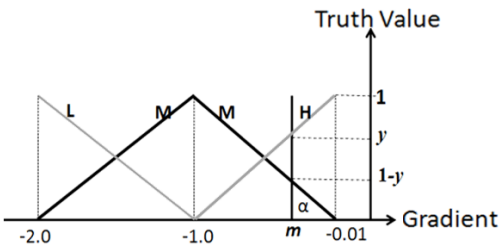

Figure 8: Fuzzy membership function of gradient and c

Once both gradient-m and intersection-c is found using the fuzzy mathematical model, the predicted equation of the motion path of the obstacle is finalized.

In order to test the theory in real world applications, a mobile robot in Figure 9 was especially prototyped. The relationships between the desired angle to be turned, encoder (Figure 10) count, the radius of robot / drive wheels was mathematically derived. And rules i.e. distance to be moved away depending upon an estimated risk factor were compiled based on experimental observations. Instructions to be executed in order to avoid obstacles one the predicted motion path is given, were stored on a Single-Chip microcontroller based Arduino Mega 2560 powered embedded system that handles the task, controlling the actuators of the model robot.

4. EVALUATION

For a sample size of 60 test cases, the implemented Fuzzy Mathematical Model possessed a maximum error for estimating the gradient of the object's motion path m as 0.089 while the maximum m error reported was 13.33 % as a percentage. The average m error reported was 0.021, meanwhile keeping an average m error as lower as 2.78 %. The result can be further summarized with an average m accuracy: 95.55 %.

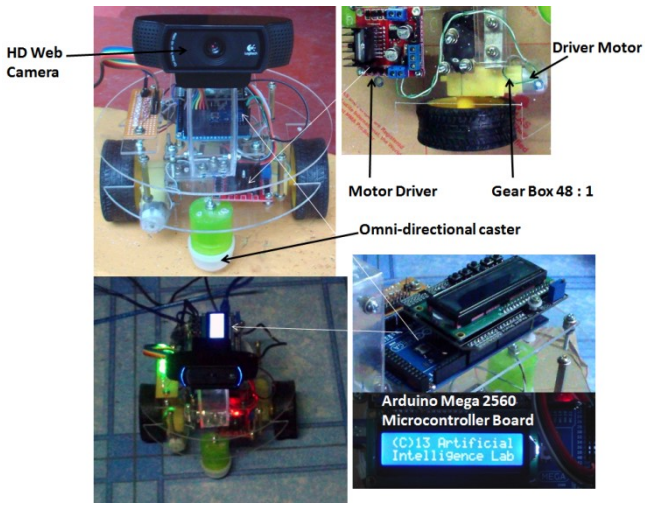

Figure 9: The mobile robot

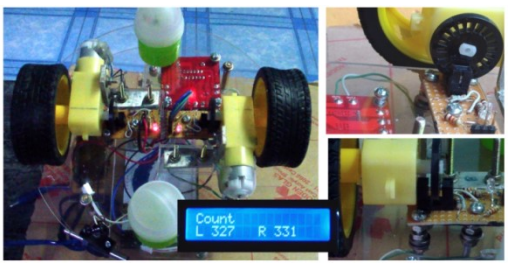

Figure 10: Velocity encoders

Prediction of intersection c by the Fuzzy-Mathematical Model is as follows:

The maximum *c* error reported was 30 (only one instance, in most cases it was below 15), while it was a maximum of 48 % as a percentage. The average *c* error reported was low as 5.12, which is 6.34 % as a percentage. (Figure 11 denotes the statistical analysis results). However it could maintain an average c accuracy 91.99% (approximately 92%) for the whole cases.

In general, if the minimum of m accuracy and c accuracy concerned, the model is precise as to

produce predictions with an average accuracy of 92%.

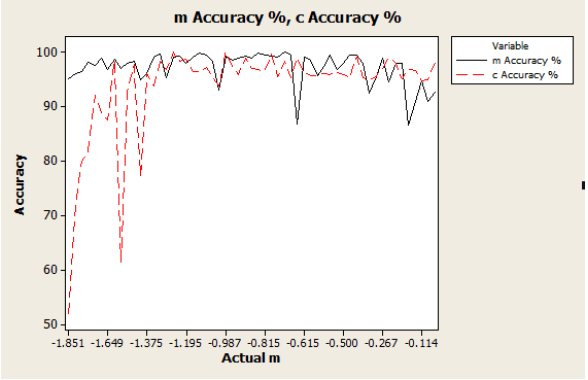

Figure 11 : Accuracy of estimated m and c, as a percentage

5. DISCUSSION AND FUTURE WORK

Experimental results and statistically analyzed evaluation reports produced with proper analytical techniques convey that the Fuzzy-Mathematical approach is appropriate to model the motion behaviour of an observed object and recover its motion path using the apparent size variation graph as the primary resource. The concept is appropriate, applicable and adaptable with real time applications. In addition, it provides benefits such as consuming a minimal computational resource, since it is an extremely nature-inspired optimized technique. Other than the less computational overhead, monocular vision is adequate instead stereo-optics where complicated mathematics and mechanisms required for calibration, align and depth estimation.

The theory postulated here can be further extended to model motion behaviours of objects in 3-D coordinate systems and having complicated and variable motion paths rather than simple linear motion.

A blind guidance system when voice instructions integrated, or a directly controlled electrically powered wheel-chair that safeguards the disable user from risky objects i.e. incoming vehicles, would be possible real world applications of the concept innovated.

REFERENCES

[1] Ashutosh Saxena, Sung H. Chung, and Andrew Y. Ng, *Learning Depth from Single Monocular Images* [Online], Available: http://ai.stanford.edu/~asaxena/learningdepth/NIPS_LearningDepth.pdf

[2] Ashutosh Saxena, Min Sun and Andrew Y. Ng, *Make3D: Depth Perception from a Single Still Image* [Online], Available: http://robotics.stanford.edu/~ang/papers/aaai08-Make3dDepthPerceptionSingleImage.pdf

[3] Francisco Bonin-Font, Alberto Ortiz, Gabriel Oliver, *Visual Navigation for Mobile Robots: a Survey,* [Online], Available: http://dmi.uib.es/~fbonin/survey.pdf

[4] Yoko Watanabe, *Stochastically Optimized Monocular Vision-Based Navigation And Guidance* [Online], Available: http://smartech.gatech.edu/jspui/bitstream/1853/22545/1/watanabe_yoko_200804_phd.pdf

[5] Animesh Garg, Anju Toor, Sahil Thakkar, Shiwangi Goel, Sachin Maheshwari, Satish Chand, *Object Identification and Mapping using Monocular Vision in an Autonomous Urban Driving System* [Online], Available: http://www.ijcte.org/icmv/ICMV2010/136-ICMV2010-W12016.pdf

[6] Tobias Low, Antoine Manzanera, *Ground-Plane Classification for Robot Navigation* [Online], Available:http://www.ensta-paristech.fr/~manzaner/Publis/icarcv10.pdf

[7] Nicolau Leal Werneck, Anna Helena Reali Costa, *Mapping with Monocular Vision in Two Dimensions* [Online],Available: http://nwerneck.sdf.org/almoxarifado/WerneckCosta10.pdf

[8] Chau Nguyen Viet, Ian Marshall, *VISION-BASED OBSTACLE AVOIDANCE FOR A SMALL, LOW-COST ROBOT* [Online], Available: http://eprints.lancs.ac.uk/27256/1/27256.pdf

[9] Jiandong Tian, Yandong Tang, *Learning and Vision-Based Obstacle Avoidance and Navigation* [Online], Available: http://ir.sia.ac.cn/bitstream/173321/9008/2/ZZZJ000014.pdf

[10] Jan Hoffmann, Matthias J¨ungel, Martin L¨otzsch,*A Vision Based System for Goal-Directed Obstacle Avoidance* [Online], Available: http://www2.informatik.hu-berlin.de/~juengel/papers/hoffmann-juengel-loetzsch-rc04.pdf

[11] Jan Hoffmann, Matthias J¨ungel, Martin L¨otzsch, *A Vision Based System for Goal-Directed Obstacle Avoidance used in the RC'03 Obstacle Avoidance Challenge* [Online], Available: http://martin-loetzsch.de/publications/rc04-obstacle.pdf

[12] Rafaelmun˜ Oz-Salinas, Eugenio Aguirre, Miguel Garci´A-Silvente, Antonio Gonza´ Lez, *Door-detection using computer vision and fuzzy logic* [Online], Available: http://www.wseas.us/e-library/conferences/athens2004-b/papers/474-224.pdf