Time-To-Contact Forecasting by Modeling the Apparent Size of Obstacles

Angel Sanchez Garcia[1], Homero Rios Figueroa[1], Gustavo Quintana Carapia[2],
Aurora Montano Rivas[3], Antonio Marin Hernandez[1]
[1]Research Center of Artificial Intelligence, University of Veracruz, 91090, Mexico
[2]ELEC Department, Vrije Universiteit Brussel, 1050, Belgium
[3]School of Statistics and Informatics, University of Veracruz, 91120, Mexico
angesanchez@uv.mx

***Abstract*—**
The growing demand for rapid response to tasks that robotic systems carried out, makes us increasingly think about robust and fast approaches. This paper presents a novel approach based on modeling the task known as Time-to-contact, in order to forecast the time it would take for a robot to collide with a detected object. For this we use an statistical approach based on time series. In addition, the location of potential obstacles is performed by passive sensors (monocular vision) in order to save energy and money than other more sophisticated sensors. Experimental results from the application of the method on real scenarios from other publications are reported.

***Keywords*—Time-To-Contact, Tau-margin, Time series, Apparent size, ARIMA model.**

I. Introduction

Robotics currently plays an important role in specific societal tasks such as assistance, home and office service, and industrial tasks that require precision or may be dangerous to human. The above activities can be divided into subtasks which themselves are not trivial, and have been handled over time [1], such as localization, communication, path planning and collision avoidance. Of these activities, this work addresses the avoidance of obstacles.

So, the mobile robots need a measure for assessing the proximity of obstacles. A measure that is well suited to this purpose is provided by the well known term called *Time-to-contact*. The Time-to-contact (TTC) is the time that would elapse before the *Center of Projection (COP)* (for example a camera of a robot) reaches the surface that it is seeing if the current relative movement between the camera and the surface remains unchanged. The Time-to-contact is usually expressed in terms of the speed and the distance of the considered obstacle. The classical equation to compute the TTC is given by (6).

$$TTC = -\frac{Z}{\frac{dZ}{dt}} \tag{1}$$

where Z is the distance between the observer and the obstacle, and $\frac{dZ}{dt}$ is the velocity of the robot with respect to the obstacle. However, with a monocular camera only, the distance is generally unknown [2].

To detect obstacles, we use computer vision because our goal is to recover from multiple images over time, the relative motion between an observer and the environment. Although vision based forward collision warning (FCW) turns out to be a rather challenging task, due to the huge amount of different environmental conditions that influence the perception of the environment, in this paper, we propose a strategy to mitigate factors that affect the perception of the world, and provide more robust and fast estimates.

To analyze our results we present two cases, reported in [3] where obstacles with different shapes (straight and convex) are used. Case I: the estimation of Time-to-contact where the obstacle is a red cylinder and the robot approaches the cylinder. Case II: The object is a green ball and the robot approaches it.

This paper is organized as follow. In section II, related work about Time-To-Contact in mobile robots is presented. The perception model of the world is shown in section III. Our proposal is provided in Section IV. In Section V the The model used to forecast new values is explained. In Section VI the details and results of the experiments are shown. Finally section VII draws conclusions.

II. Related Work

The Time-to-contact is not new. In fact, Time-to-contact estimation itself was first studied and defined in [4] as the distance to an obstacle divided by the relative velocity between them. Over time approaches have been developed to estimate it more accurately, however, there are still present errors with respect to ground true because each environment is different. These approaches are classified in the diagram of Fig. 1, and explained below. The red box indicates the place where our work is located.

First, we mention that in the area of autonomous mobile robots, specifically in the navigation task, it is common to provide them with specialized sensors such as sonar [5][6], binocular vision (stereo vision) [7] where it has two generally calibrated cameras looking at a point, paracatadioptric sensors with good results in real time [8], ultrasonic range sensors, and infrared sensors which detect a smaller field of view have been used. However, the fact of having other kind of sensors often expensive, and it involves an increase in energy expenditure used by the robot, so we decided to use approaches where energy and money saving is maximized and therefore we have

978-1-5386-2532-3/17 $31.00 © 2017 IEEE
DOI 10.1109/ICMEAE.2017.12

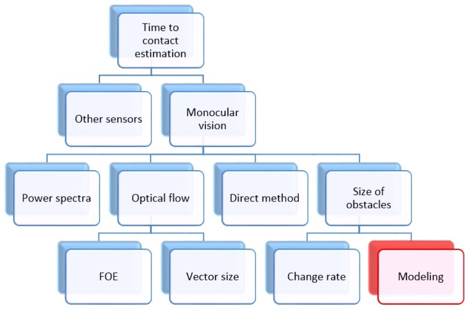


Figure 1: Classification of Time-to-contact approaches.

paid special attention in methods by one camera (monocular vision).

Other kind of approach is to estimate Time-to-contact from images in space domain or in spectral domain [9] (using temporal change of power spectra between successive images. Many approaches are based on optical flow calculation [10][11][12] because is possible to estimate the Time-to-contact having access to the translational component of the flow field. From optical flow, it is possible to estimate the focus of expansion (FOE) and estimate the Time-to-contact on not flat surfaces [13][14]. Having calculated the flow field, it is also possible to detect relative closeness of objects, according to the length of optical flow vectors as in the approaches shown in [15][16]. However, since these approaches are based on estimating optical flow, they are iterative, need to work at multiple scales, tend to be computationally expensive and require a significant effort to implement properly [17].

Another kind approach is called "Direct method". In fact, it is so named because it works directly with the derivatives of image brightness and does not require feature detecting, feature tracking, or estimation of the optical flow [18]. Another approach is based on the change of size of obstacles. In works as [2][19][3], the authors are based on the fact that animals and insects obtain information from the apparent size s of objects and the temporal changes in the size $\frac{s}{\frac{ds}{dt}}$. This information is called the"tau-margin", which is is derived from (6) by using a characteristic size of the obstacle in the image [20] is defined as (4).

$$\tau = -\frac{s}{\frac{ds}{dt}} \tag{2}$$

Finally, another approach is based on modelling the movement of the robot. In [21] the authors take three cases: The first model assumes a constant, non-zero relative acceleration between the obstacle and the host car. The second model assumes zero acceleration but non-zero relative velocity and the third model assumes a constant distance. In that method, the obstacles are found using interest points with a Harris corner detector. The idea is good, however the use of interest points leads to that we can only have one object in the scene. If there are more objects all points will be part of a single object (region). For the above features, advantages and disadvantages of the each kind of approach, we decided to place our work as is shown in the red box of Fig. 1.

III. Time-To-Contact estimation

Animals obtain information of time-to-contact from the apparent size of objects and temporal changes in the size. In [19] is stating that animals get the $tau - margin$ from temporal changes of apparent size of objects, Fig. 2 shows the mechanism of visual perception of an object, where S is the size of the object and r denotes the distance between the object and the observation point. The rate of change of apparent size over time, is given by (3).

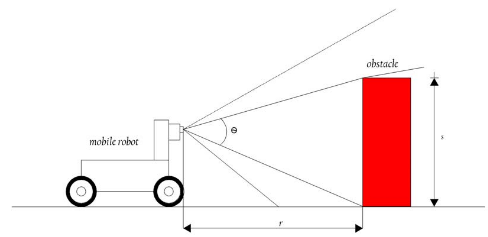


Figure 2: Visual perception of an object.

$$\frac{d\theta}{dt} = -\frac{S}{r^2}\frac{dr}{dt} \tag{3}$$

Equation (4) reveals how the $tau - margin$ is calculated and how it equals the calculate of change in distance r per unit time. It also shows that τ can be determined directly from the θ and its temporary change, without requiring information on the distance or the relative speed of the object.

$$\tau = -\frac{\theta}{\frac{d\theta}{dt}} = \frac{r}{\frac{dr}{dt}} \tag{4}$$

In the segmentation process, the robot extracts information from the scene with a color segmentation process. The apparent size of the obstacle is measured as an aperture angle whose vertex is in the center of the camera lens as is defined in [3] and specified in (5) where f is the focal length. The larger the aperture angle, the nearer the obstacle is to the robot.

$$\theta = 2\left[\arctan\left(\frac{\frac{S'}{2}}{f}\right)\right] \tag{5}$$

For *Tau-margin* τ, we need to see the change in apparent size as is shown in (6). For this, we have that the variation of θ with respect to time taking the variation as 1 frame, and so, τ can be calculated from equation (7).

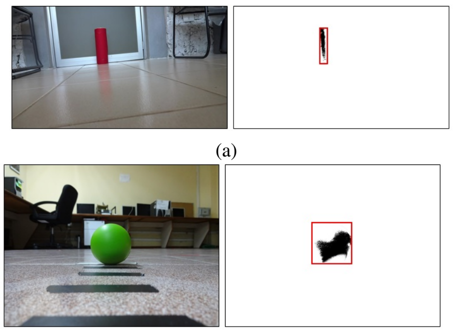
$$\dot{\theta} = \frac{d\theta}{dt} \approx \frac{\theta_t - \theta_{t-1}}{\Delta = 1} \tag{6}$$

$$\tau_t = -\frac{\theta_t}{\dot{\theta}_t} \tag{7}$$

The units of τ depend of the time slot with which the image was taken. That is, if each frame is analyzed, units of τ will be frames.

IV. Our Proposal

As mentioned above, we aim to estimate the Time-to-contact is as robust and fast as possible, therefore, we base our work on modeling the behavior of the phenomena of approaching an obstacle with constant velocity.



(a)

(b)

Figure 3: Examples of segmentation by color. a) Case I: shows the segmentation of an obstacle by red color in a real image. b) Case II: shows the segmentation of an obstacle by green color in a real image.

As we might intuit, while the robot approaches, the change of apparent object size increases. As for the first requirement (robustness), as is reported in [3], it can be possible to have values that they do not have a similar behavior to the other values (outliers) due to different factors, such as change in light intensity, unstable movement of the robot because the surface on which it moves is not flat or simply by noise in the sensor (camera). These changes cause the $tau-margin$ is positive when the apparent size of the object at time t is greater than at time $t-1$ and negative when the apparent size of the object is less at time t than at time $t-1$, and therefore, the Time-to-contact is not reliable. So, we expect that in modeling, we are describing the general growth behavior of the size of the segmented obstacle and we could have a more stable Time-to-contact estimation.

On the other hand, modeling behavior based on the detection of some frames and predicting those measurements in the future, will help to the robot to process its task faster because it will not have to be looking the obstacle in each frame. It will save time processing once the model is created.

To choose the model that best fits the measurements or input signal, seen as a time series, we start with the idea that the robot is approaching a detected obstacle, and therefore, if we plot the measurements of the apparent size, we will find that it is an increasing series (the series has a positive tendency) and therefore it would not be stationary in the mean, since it would have an infinite mean. Consequently, in order to adjust a model, the trend must be removed and a differentiation operation must be carried out as a first step.

We performed the Phillips-Perron test [22], based on Dickey-Fuller test, to the original series and to the differentiated series, in order to corroborate that when transforming the series, these become stationary, and also that only a parameter of differentiation is needed. Table 1 shows the results of applying this test before and after differentiation, where it can be observed that with a differentiation, $p-value$ is less than .05, so there is sufficient statistical evidence to reject the null hypothesis of the test, and these results indicate that the series with first order differentiation, are stationary.

TABLE I: Results of Phillips-Perron test

Case	Dickey-Fuller Statistic	p-value
Case 1 Original	-4.5747	0.8535
Case 2 Original	-12.4707	0.4056
Case 1 Differing	-283.8433	0.01
Case 2 Differing	-256.0387	0.01

With the above, we decided to treat the estimates with an ARIMA model, since it includes parameters of differentiation to work with a stationary series. We do not model the TTC as such, but we model the behavior of the apparent size of the projected obstacle in the image because the Time-to-contact depends on this growth.

V. ARIMA model

Time series forecasting is the use of a model to predict future values based on previously observed values. ARIMA *Autoregressive Integrated and Moving Average* is a probabilistic model which assumes that errors have a normal distribution with mean zero and variance σ^2. This assumption is called *White Noise*. Also, this assumes that there is no autocorrelation in the errors. The expression used to represent this assumption is shown in (8).

$$\varepsilon_t \sim N(0, \sigma^2) \tag{8}$$

While exponential smoothing methods do not make any assumptions about correlations between successive values of the time series, in some cases we could make a better predictive model by taking correlations in the data into account.

The autoregressive (AR) model [23][24] specifies that the output variable depends linearly on its own previous values. Its set-up is based on using data observed in the past to develop the AR model coefficients. *Moving Average* is one of widely known technical indicator used to predict the future data in time series analysis [25]. MA is a common average of the previous n data points in time series data. Each point in the time series data is equally weighted.

The autoregressive and moving average process, denoted as $ARMA(p,q)$, is a combination of an autoregressive process of order p and moving average process of order q. This combination can be written as it is shown in the equation (9):

$$\begin{aligned} X_t = c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + ... + \phi_p X_{t-p} \\ + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + ... + \theta_q \varepsilon_{t-q} \end{aligned} \tag{9}$$

Where ε_t errors satisfy a white noise sequence. The ARIMA model symbolized as *ARIMA(p,d,q)*, follows a similar process that *ARMA(p,q)* processes, only now in the initial model is taken into account that the series is not stationary and it is possible that differentiations of some order are needed.

Therefore p denotes the number of autoregressive model parameters, d denotes the order of the differentiations required to solve problems of non-stationarity and q denotes the number of parameters required moving averages. This kind of model is crucial, because it helps us to engage the *trend* and *periodicity* in the time series to the model, if these elements are presented. To select the values of p and q, several models are generated and tested to select the adjusted model with lower values of Akaike information criterion AIC (which is a measure of the relative quality of a statistical model for a given set of data) and σ^2.

VI. Experiments and results

For the experiments described in this paper, we took as example videos of 210 frames reported in [3] where they were taken with a frequency of 30 frames per second. The size of each image was 1920 x 1080 pixels. The colors used to segment the obstacles were for red R = 210, G = 50 and B = 50 and for green R = 20 G = 220 and B = 20 with a threshold = 50, which were obtained experimentally. Fig. 4 shows the scenario to estimate the Time-to-contact.

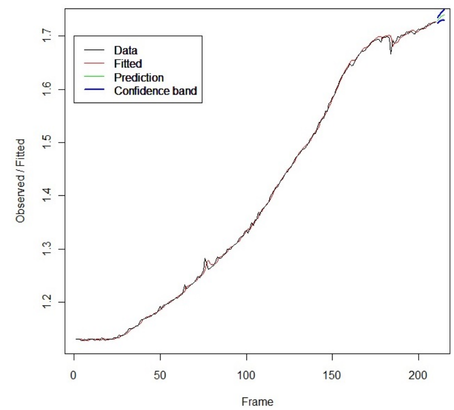
Figure 4: Example of scenario of the experiments.

In Fig. 5 is shown in black the data of each periodogram. The fitted model is shown in red. In green color the forecast for the next 5 frames is drawn and in blue the confidence band is shown, indicating the variability of the forecasts. These lines representing the upper and lower confidence limits for all points of a line fitted within the range of the data. Fig. 5 shows that the fitted model that describes the behavior of the apparent size change of the obstacle over time, so we will conclude that the forecast statistically reliable.

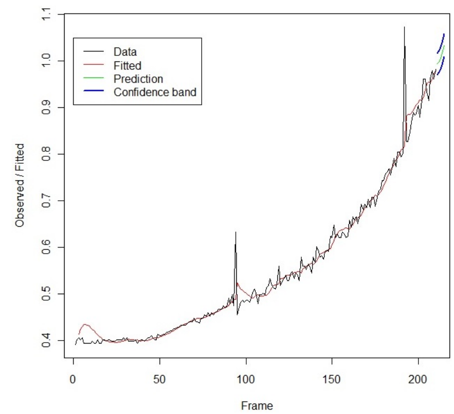
A way to validate the models is to calculate the sum of squared errors (SSE), which is a measure of the accuracy of the forecasts, that is, the forecast errors for the time period covered by our original time series. The forecast errors are calculated as the observed values minus predicted values, for each time point. Table II shows the estimated SSE for each model and it can be seen that they are really small, which makes us think that the models are fitted in such a way that the forecasts are reliable.

TABLE II: sum of squared errors SSE

Case	SSE Statistic
Case 1 Original	0.1400
Case 2 Original	0.0045



(a)



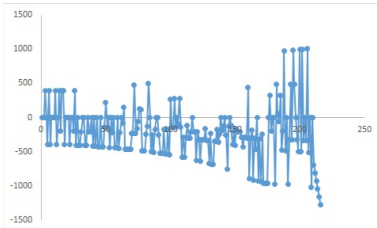
(b)

Figure 5: Fitting the model to the data and forecast. a) Case I: Forecasting of Apparent size for the perception of the red cylinder. b) Case II: Forecasting of Apparent size for the perception of the red cylinder.

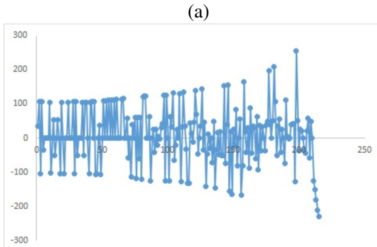
In Figure 6, the Time-to-contact estimations are observed. From Frame 210 we show 5 forecasts with our method. As can be seen in this figure, due to situations of perception of the environment, the observations of the apparent size are not smooth (ie they do not grow uniformly) and therefore the TTC estimates are unreliable and they have erratic behavior. However, when we model the behavior, the TTC estimate stabilizes.

VII. Conclusions

In this paper, a novel method to forecast the Time-to-contact is presented. To identify the obstacle, segmentation by color is proposed, and to identify apparent size of the object was shown that this is estimated by the height or width, without using growth of regions. We also exemplify the change in apparent size of obstacles to calculate the $tau - margin$ in previously reported real scenarios. Models and obtained data were compared and statistically analyzed visually and quantitatively. The main contribution consists of predicting the growth of the obstacle detected to directly estimate time it would take the robot to collide with an obstacle instead of



(a)



(b)

Figure 6: TTC estimations modeling with 5 frames. a) Case I: shows the TTC estimation of the red cylinder. b) Case II: shows the TTC estimation of the green sphere.

calculating it in each frame. With this we are saving processing time and cost by using passive sensors like a camera. When generating a model of how the apparent size change of a detected obstacle is behaving, we discard measures that do not have a behavior equal to the rest, and that can be caused by factors of the environment or some problem in the sensor, with which we are strengthening our estimates. As future work we will work on better segmenting obstacles to not depend on a priori knowledge of them, such as shape or color. And therefore more than one obstacle can be detected in the scene.

Acknowledgment

The first author thanks to CONACYT for its financial support and University of Veracruz to the development of this part of PhD work.

References

[1] A. Abdulla, H. Liu, N. Stoll and N. Thurow, *A New Robust Method for Mobile Robot Multifloor Navigation in Distributed Life Science Laboratories*, Journal of Control Science and Engineering, Vol. 2016, pp. 1 - 17, 2016.
[2] G. Alenya, N. Negre and J. Crowley, *Time To Contact for Obstacle Avoidance*, 4th European Conference on Mobile Robots ECMR 09, pp. 19 - 24, 2009.
[3] A. Sanchez, H. Rios, A. Marin, K. Cortes, G. Contreras, *Estimation of Time-To-Contact from Tau-margin and Statistical Analysis of Behavior*, 2016 IEEE International conference on Systems, Signals and Image Processing (IWSSIP), pp. 1-6, 2016.
[4] D. N. Lee, *A theory of visual control of braking based on information about time-to-collision*, Perception, Vol 5, pp. 437-459, 1976.
[5] B. K. Shim, J. S. Yang, E. G. Kim, Y. K. Jeong, J.B. Won, S. H. Han, *A travelling control of mobile robot based on sonar sensors*, 15th International Conference on Control, Automation and Systems (ICCAS), pp. 1241 - 1243, 2015.
[6] J. Vascak and J. Hvizdos, *Vehicle navigation by fuzzy cognitive maps using sonar and RFID technologies*, 14th International Symposium on Applied Machine Intelligence and Informatics (SAMI), pp. 1 - 6, 2016.
[7] M. Muffert, T. Milbich, D. Pfeiffer and U. Franke, *May I enter the roundabout? A time-to-contact computation based on stereo-vision*, IEEE Intelligent Vehicles Symposium (IV), pp. 565 - 570, 2012.
[8] F. Z. Benamar, S. E. Fkihi, C. Demonceaux, E. Mouaddib and D. Aboutajdine, *Gradient-Based time to contact on paracatadioptric camera*, 2013 International Conference on Image Processing, ICIP'2013, pp. 5 - 9, 2013.
[9] W. Yu, G. Sommer and K. Daniilidis, *Multiple motion analysis: in spatial or in spectral domain?*, Computer Vision and Image Understanding, Vol. 90 pp. 129 - 152, 2003.
[10] B. Horn and B. Schunk *Determining optical flow*, Journal of Control Science and Engineering, Vol. 17, pp. 185 - 203, 1981.
[11] F. Meyer, *Time-to-collision from first-order models of the motion field*, Journal of Control Science and Engineering, Vol. 10, pp. 792 - 798, 1994.
[12] F. Meyer and P. Bouthemy, *Estimation of time-to-collision maps from first order motion models and normal flows*, Proceedings of the 11th IAPR International Conference on Pattern Recognition, pp. 78 - 82, 1992.
[13] C. McCarthy, N. Barnes and R. Mahony, *A Robust Docking Strategy for a Mobile Robot Using Flow Field Divergence*, IEEE Transactions on Robotics, Vol. 24, pp. 832 - 842, 2008.
[14] C. McCarthy and N. Barnes *A Unified Strategy for Landing and Docking Using Spherical Flow Divergence*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 34, pp. 1024 - 1031, 2012.
[15] Y. S Lau, Q. Zhang, Y. Li and S. S. Ge, *Non-Metric Navigation for Mobile Robot Using Optical Flow*, IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 4953 - 4958, 2012.
[16] A. Sanchez, H. Rios, A. Marin and G. Contreras, *Decision making for obstacle avoidance in autonomous mobile robots by time to contact and optical flow*, International Conference on Electronics, Communications and Computers, pp. 130- 134, 2015.
[17] B. K. P. Horn, Y. Fang and I. Masaki, *Hierarchical framework for direct gradient-based time-to-contact estimation*, Intelligent Vehicle Symposium, pp. 1394 - 1400, 2009.
[18] B. K. P. Horn, Y. Fang and I. Masaki, *Time to Contact Relative to a Planar Surface*, Proceedings of the 2007 Intelligent Vehicles Symposium, pp. 68 - 74, 2007.
[19] Y. Kaneta, Y. Hagisaka and K. Ito *Determination of Time to Contact and application to timing control of mobile robot*, Proceedings of the 2010 IEEE, International Conference on Robotics and Biomimetics 2010.
[20] A. Negre, C. Braillon, J. L. Crowley and C. Laguier *Real-time time-to-collision from variation of intrinsic scale*, Proceedings of the 10th International Symposium on Experimental Robotics: pp. 75 - 84, 2006.
[21] D. Muller, J. Pauli, C. Nunn, S. Gormer and S. Muller, *Time To Contact Estimation Using Interest Points*, Proceedings of the 12th International IEEE Conference on Intelligent Transportation Systems, 2009.
[22] P. Phillps, P. Perron, *Testing for a unit root in time series regression*, biometrika, vol. 7, 2, pp. 335-346, 1988.
[23] S. Haykin, *Topics in Applied Physics, Nonlinear Method of Spectral Analysis*, Vol. 34 Springer-Verlag, 1983.
[24] S. M. Kay and S. L. Marple, *Spectrum Analysis- a modern perspective*, Proc. IEEE, vol. 69, pp. 1380 - 1419, 1981.
[25] S. Hansun, *A New Approach of Moving Average Method in Time Series Analysis*, Microwave Symposium Digest, New Media Studies (CoNMedia), 2013.