Motion Prediction of Regions Through the Statistical Temporal Analysis Using an AutoRegressive Moving Average (ARMA) Model

Angel Juan Sánchez García[1], Homero Vladimir Ríos Figueroa[2], María de Lourdes Velasco Vásquez[1], Gerardo Contreras Vega[1], Antonio Marín Hernández[2]

[1] Faculty of Statistics and Informatics, University of Veracruz, Xalapa, Veracruz, México
[2] Department of Artificial Inteligence, University of Veracruz, Xalapa, Veracruz, México

{angesanchez, hrios, lovelasco, gcontreras, anmarin}@uv.mx

Abstract. Currently many applications require tracking moving objects, and that information is used to plan the path of motion or change according to the position of a visual target [1] [2]. Computer vision systems can predict the motion of objects if the movement behavior is analyzed over time, ie, it is possible to find out future values based on previously observed values. In this paper, a proposal to predict motion of segmented regions is presented, through an analysis of a time series using an ARMA model. Two scenarios with different characteristics are presented as test cases. Segmentation of moving objects is done through the clustering of optical flow vectors for similarity, which are obtained by Pyramid Lucas and Kanade algorithm.

Keywords: Prediction, motion, optical flow, time series, ARMA.

1 Introduction

Computer vision has become a tool that provides us the information for "smart" interaction with the environment without being physically in contact with it [3]. One of the primary tasks of computer vision is to reconstruct, from two-dimensional images, the 3D properties of a scene, such as shape, motion and spatial order of objects. In monocular vision, an important goal is to recover from multiple images over time, the relative motion between an observer and the environment. The structure of the environment obtained is used to generate relative distances between points on a surface in the scene and the observer [4].

Many applications that incorporate computer vision are used to track moving objects. To understand the motion of objects in a series of frames, there are two main tasks: identification and description [5]. Identification is finding the objects of interest in a set of frames from a video stream. The second task would be to describe the movement. With the analysis of motion over time, we can create a model that allows us to predict, with some confidence, the position in the following frames.

The outcome is the result of the analysis of phenomena or events that occur over time. We propose the application of an ARMA (autoregressive moving average) model to predict, based on the temporary records, the position of moving objects. We use two scenarios as test cases. In the scenario 1 is shown the prediction of vehicle moving from left to right. In the second scenario is show an arm with a chess board moving randomly.

This paper is organized as follow. In section 2, related work about tracking motion and ARMA models are presented. A brief description of our proposal is shown in section 3. The methodology for the segmentation of moving objects in regions is provided in Section 4. In Section 5 the description of the technique to build a model of motion is explained. In Section 6 the details of the experiments and the results are shown. Finally section 7 draws some conclusions and future work.

2 Related Work

For decades, it has attempted to identify moving objects, track them, and to anticipate their movement. In recent years, optical flow has been used for segmenting moving objects. In [6] the authors show results of an application to segment moving cars from points of interest formed by optical flow vectors obtained by the method of pyramid Lukas-Kanade. Since in this application is only for cars, it only detect moving objects from a priori information on the object type. Moving objects are detected from the image regions with nonzero optical flow vectors and grouped according to the moving speed and Euclidean distance. The Segmentation of the detected objects is performed using a priori information of the same shape, so that segmentation is carried into rectangles of similar size. Once defined the regions that are in motion in the image sequence, we could predict the motion of regions in consecutive frames.

A common technique for tracking and motion prediction is the Kalman Filter. This method has been used in several areas when it is looking to describe the motion of objects and that somehow it could be found or measured the position of these. In [7] the Kalman filter was used to generate a model that it could predict the movement of regions. This model has two phases: prediction and adjustment. In this last phase, the parameters of the model are adjusted depending on the error caused in the prediction phase. However, other approaches have been used to predict the motion in sequences of images.

In [8] is proposed a model for the interframe correspondences existing between pixels of an image sequence. These correspondences form the elements of a field called the motion field. In their model, spatial neighborhoods of motion elements are related based on a generalization of autoregressive (AR) modeling of time-series. Also in [9] a framework for predicting future positions and orientation of moving obstacles in a time-varying environment using autoregressive model (AR) is described. The AR model has been used in other fields, because it has a low complexity of computation [10], for example in the motion of vechicles [10][11], or in the medical field to predict the evolution of a tumor [12][13]. Even the AR model can be used to modify the model performed in Kalman filter [14]. However, using AR model does not involve to get parameters of the trend in a time series. On the other hand, In [15] an aggregation approach is proposed for traffic flow prediction that is based on the moving average (MA), exponential smoothing (ES), auto-regressive MA (ARIMA), and neural network (NN) models. However the aggregation approach assembles information only from relevant time series of the traffic flow volume that it is collected 24 h/day over several years.

3 Our Proposal

Because we can know the position of the region (object) in time, it is possible to predict the next position from these records. A time series is a set of chronological observations of a phenomenon that it is occurring in a defined time period. So, we want a model that incorporates the previous records of that phenomenon and it could be satisfied with an Auto Regressive model (AR), but also the model must involve trend and periodicity (if these elements are present in the behavior) and we can use a Moving Average model (MA). So we propose using a combined model with both elements (ARMA) to predict the motion of segmented regions by optical flow.

4 Segmentation

The ability to detect motion is crucial for vision and guided behavior by the sense of sight [16]. So, if we want to identify objects that are moving, we need to have at least two images. This is because it is necessary to know if there was any change in the intensity of the pixels between the two images, allowing us to identify a motion. Object tracking systems require accurate segmentation of the objects from the background for effective tracking [17]. To identify object that have moved from one scene to another, the optical flow is often used.

For the identification and segmentation of moving objects in regions, it is used the methodology proposed in [18], which consists in to select the points of interest in the image where the optical flow can be reliable. Harris called to these points of interest corners [19] and they are obtained by the method of Shi and Tomassi [20], where the eigenvalues λ_1 and λ_2 are calculated from the autocorrelation matrix of each pixel, and we can obtain one of three possible cases:

1. If $\lambda_1 \approx 0$ and $\lambda_2 \approx 0$, then the pixel does not have relevant information.
2. If $\lambda_1 \approx 0$ and λ_2 has a large positive value, then it is an edge.
3. If λ_1 and λ_2 are large positive values, then the pixel is a corner.

If the minimum eigenvalue exceeds a threshold α, then the corner has a reliable value to be identified and it is a good point to track. Later the optical flow is calculated by the method of Lucas and Kanade [21] only at points of interest. In the figure 1 are shown examples of the optical flow vectors in both scenarios.

After obtaining the optical flow vectors, they are grouped by similarity. To create each region that define each object, it is necessary to group the optical flow vectors based on 3 criteria: proximity, direction and magnitude.

Then, for each group of vectors, their convex hull is obtained to address the problem of discontinuity of optical flow regions and segment the region [18]. The points used to generate the convex hull of each region are the starting points of the vectors belonging to the region.

(a) scenario 1 (b) scenario 2

Fig. 1. Optical flow vectors from Lucas and Kanade Method.

We can represent the condition to determine whether two points belong to the convex hull from the set of points *S*. Two points **P**(x_1, y_1) and **Q**(x_2, y_2) belong to the set of the covex hull *C* if and only if all points **R**(x_3, y_3) belonging to *S* (except **P** and **Q**) when they are evaluated in the equation of the line through the points **P** and **Q**, are on one side of the line (the sign must have a single value for every point **R**, either positive or negative). In other words, Equation 1 is satisfied or Equation 2 is satisfied, but not both.

$$\boldsymbol{P}(x_1,y_1) \in C \text{ y } \boldsymbol{Q}(x_2,y_2) \in C \leftrightarrow \forall \boldsymbol{R}(x_3,y_3) \in S \mid (y_2 - y_1)x_3 - (x_2 - x_1)y_3 < (y_2 - y_1)x_1 - (x_2 - x_1)y_1,\ \boldsymbol{P} \neq \boldsymbol{Q} \neq \boldsymbol{R}. \quad (1)$$

$$\boldsymbol{P}(x_1,y_1) \in C \text{ y } \boldsymbol{Q}(x_2,y_2) \in C \leftrightarrow \forall \boldsymbol{R}(x_3,y_3) \in S \mid (y_2 - y_1)x_3 - (x_2 - x_1)y_3 > (y_2 - y_1)x_1 - (x_2 - x_1)y_1,\ \boldsymbol{P} \neq \boldsymbol{Q} \neq \boldsymbol{R}. \quad (2)$$

Since the phase of segmentation, a polygon is obtained, and the polygon could change in shape and size of a pair of frames to others due various factors, so it is necessary to interact with a representative point of the polygon in the modeling process. Therefore, the centroid of each polygon is taken as reference. Figure 2 shows examples of the segmentation procedure for each scenario and their centroids.

(a) scenario 1 (b) scenario 2

Fig. 2. Segmentation of moving objects.

5 ARMA Model

Time series forecasting is the use of a model to predict future values based on previously observed values. ARMA 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 (3):

$$\varepsilon_t = N(0, \sigma^2) \ . \qquad (3)$$

5.1 AutoRegressive Model (AR)

The autoregressive (AR) model [22] [23] 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. Once the model is established, future realizations can be predicted by present occurrences [24]. The autoregressive model of order p, is expressed by (4):

$$X_t = c + \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \dots + \varphi_p X_{t-p} + \varepsilon_t \ . \qquad (4)$$

Where ε_t errors satisfy a white noise sequence and p is the number of autoregressive parameters.

5.2 Moving Average (MA)

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 moving average process of order *q*, is expressed by (5):

$$X_t = c + \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \ldots + \Phi_q X_{t-q} + \varepsilon_t \; . \qquad \textbf{(5)}$$

Where ε_t errors satisfy a white noise sequence and *q* is the number of moving average parameters. 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.

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 (6):

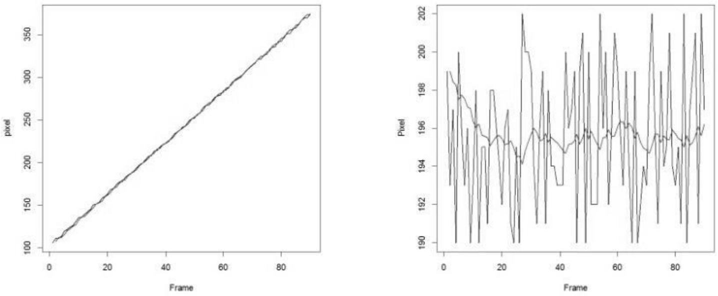
$$X_t = c + \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \ldots + \varphi_p X_{t-p} + \varepsilon_t + \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \ldots + \Phi_q X_{t-q} \; . \qquad \textbf{(6)}$$

In the combined models *ARMA(p, q)*, we have to select the values of *p* and *q*. For this, it is necessary to know how to determine the maximum values of *p* and *q* which are generated and tested in several models. To select the best model we choose 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.

6 Experiments and Results

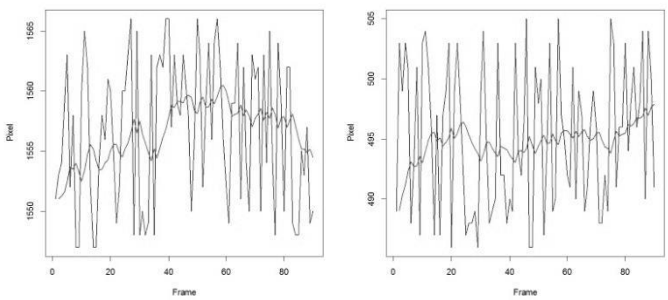
For the experiments, in the scenario 1 were used images of 640 x 480 pixels and in the scenario 2 where used images of 1920 x 1080 pixels. In both cases the images were taken from a video with a frequency of 30 images per second. The value for was $\lambda = 0.01$ and the results were obtained offline. Although the images were taken in RGB format, the image processing was performed on a single channel, i.e. gray scale.

For each block of frames, two time series are made, one of these is to describe the motion in the axis *x*, and another is to describe the axis *y*. The time series are smoothed, obtaining the ARMA model. Figure 3 and 4 show examples of adjusted model, where in blue is drawn the original time series and in red color are drawn the fitted values from the ARMA model. As can be appreciate, the fitted values have the similar behavior than the original series, but it eliminates the peaks. In figure 3, can be seen that the time series presents trend in the axis x, because the motion is from left to right and the value of the pixels in axis *x* is increasing.



(a) scenario 1, axis x (b) scenario 1, axis y

Fig. 3. Example of smoothing of the time series for scenario 1.

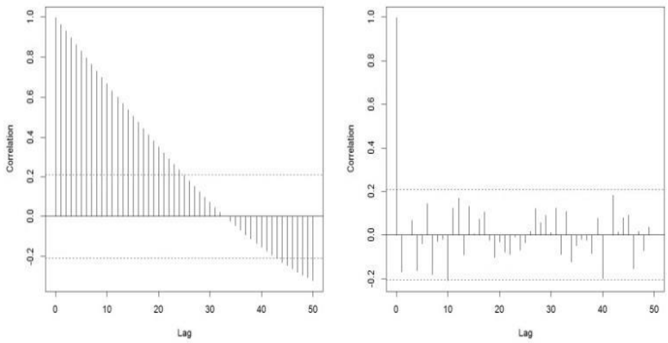


(a) scenario 2, axis x (b) scenario 2, axis y

Fig. 4. Example of smoothing of the time series for scenario 2.

Before using the ARMA model, the time series should be stationary in mean and variance, so they are all tested with the Phillips-Perron test. To find the number of tentative autoregressive parameters, the partial correlogram is calculated and to find the number of moving average parameters, the correlogram of the time series is obtained.

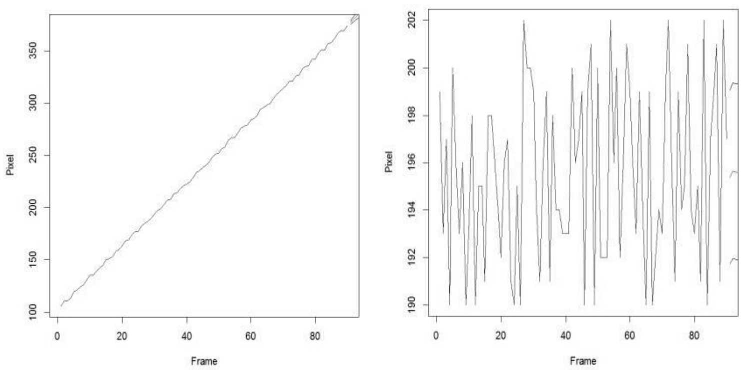
In figure 5 are shown examples of correlograms of scenario 1 with a lag of 50. In figure 5 (a) can be appreciated that there are significant correlations and they rise and fall slowly, which suggests that the time series has trend and we can test models with 24 parameters of moving average (there are 24 significant autocorrelations. Conversely, in figure 5 (b) is drawn the correlations at the same scenario 1 in axis y and the program can assume that the model does not require moving average parameters because there are not significant autocorrelations.



(a) scenario 2, axis *x* (b) scenario 2, axis *y*

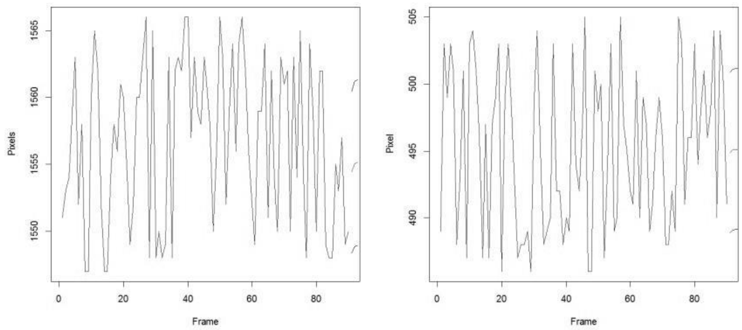
Fig. 5. Example of correlogram in the scenario 1.

The number of autoregressive and moving average parameters chosen depends of the model that has a lower value of Akaike in the combination of models. In figure 6 and 7 the predictions are presented. In red color is drawn the outcome, and the black color shows the confidence band of the model.



(a) scenario 1, axis *x* (b) scenario 1, axis *y*

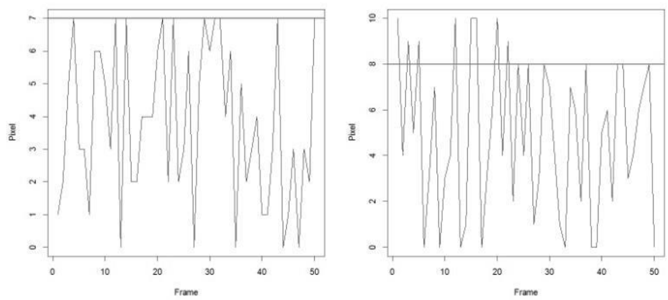
Fig. 6. Prediction using ARMA in scenario 1.



(a) scenario 2, axis *x* (b) scenario 2, axis *y*

Fig. 7. Prediction using ARMA in scenario 2.

With this method, the confidence bands presented an average of 7.5 pixels of certain in scenario 1 and 9 pixels in the scenario 2. In figures 8 and 9, are shown the errors in the outcome. The red line represents the average of significant band along 50 frames, and it can be seen that errors are rarely above the allowed limit that mark the confidence bands, and when this happens, the difference is a few pixels.

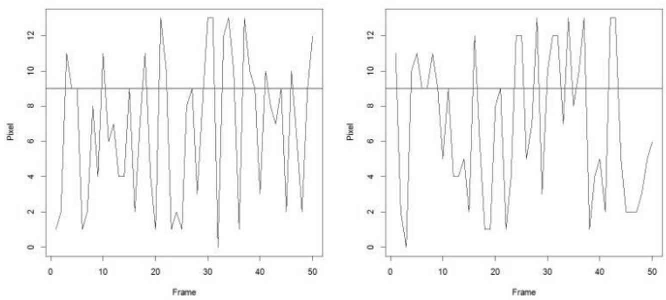


(a) scenario 1, axis *x* (b) scenario 1, axis *y*

Fig. 8. Magnitude of the errors in scenario 1.

7 Conclusions and Future Work

A proposal to predict the motion of segmented regions using an ARMA model is presented. The confidence bands present a very narrow confidence compared to the size of the captured images, similar case in the average of errors. With this process, we can involve in the model previously observed values, trend and cycles, therefore, it is possible to predict the movement of the regions (centroids) although the time series have virtually any behavior. The outcome depends of the segmentation process, but this process does not affect drastically the model, because the statistical analysis of time series gives us a confidence band to mitigate error in segmentation. This error could be caused by factors that influence in the generation of the polygon, as partial occlusion of the object, changes in light intensity or speed. As future work it can be evaluated this kind of probabilistic models with other predictive models as the Kalman filter in the same scenarios, to observe the behavior of both processes.



(a) scenario 2, axis *x* (b) scenario 2, axis *y*

Fig. 9. Magnitude of the erros in scenario 2.

References

1. Baldonado, M., Chang, C.-C.K., Gravano, L., Paepcke, A.: The Stanford Digital Library Metadata Architecture. Int. J. Digit. Libr. 1, 108–121 (1997)
2. Bruce, K.B., Cardelli, L., Pierce, B.C.: Comparing Object Encodings. In: Abadi, M., Ito, T. (eds.): Theoretical Aspects of Computer Software. Lecture Notes in Computer Science, Vol. 1281. Springer-Verlag, Berlin Heidelberg New York, 415–438 (1997)
3. van Leeuwen, J. (ed.): Computer Science Today. Recent Trends and Developments. Lecture Notes in Computer Science, Vol. 1000. Springer-Verlag, Berlin Heidelberg New York (1995)
4. Michalewicz, Z.: Genetic Algorithms + Data Structures = Evolution Programs. 3rd edn. Springer-Verlag, Berlin Heidelberg New York (1996)
1. Asano, Y. and Kawamura, A.: Decoupled rotational motion control for visual walking stabilization, IEEE Int. Workshop on Advance Motion Control (AMC 2008), pp.68-73 (2008)
2. Kim, J., Park, I. Lee, J. and Oh, J.: Experiments of vision guided walking of humanoid robot, KHR-2, IEEE-RAS Int. Conf. on Humanoid Robots, pp.135-140 (2005)
3. Horn, P.: Robot Vision, The MIT press, ISBN: 0-262-08159-8, USA (1993)
4. Negahdaripour, S.: A direct Method for locating the Focus of Expansion, Tecnical Report. A. I. Memo No 939. (1987)
5. Bradski, B. and Kaebler, A.: Learning OpenCV. Computer vision with the OpenCV library, Oreilly, First Edition, (2008)
6. D. Mora, A. Paez and J. Quiroga Deteccion de objetos Moviles en una Escena utilizando Flujo optico, XIV Simposio de tratamiento de señales, imágenes y vision artificial, STSIVA, (2009)

7. Sanchez, A., Rios, H., Marin, A. and Acosta, H.: Tracking and Prediction of Motion of Segmented Regions Using the Kalman Filter. CONIELECOMP 2014, At 88 - 93, Volume: IEEE, Proceedings, ISBN:978-1-4799-3468-3 (2014).
8. Rajagopalan, R., Orchard, M. T. and Brandt, R. D.: Motion Field Modeling for Video Sequences, IEEE Transactions on image processing, Vol. 6, No. 11 (1997).
9. Elnagar, A., Gupta, K.: Motion Prediction of Moving Objects Based on Autoregressive Model, Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE, Transactions on Volume: 28, Issue: 6 (1998).
10. Wei, D., Ye, J., Wu, X., Liang, L.: Time Series Prediction for Generalized Heave Displacement of a Shipborne Helicopter Platform, ISECS International Colloquium on Computing, Communication, Control, and Management (2008).
11. Jun, L., Sumei, W., Ke, P., Jun, X., Yun, W., Tao, Z.: Research on On-line Measurement and Prediction for Vehicle Motion State, International Conference on Digital Manufacturing and Automation (2010).
12. Ichiji, K., Sakai, M., Homma, N., Takai, Y. and Yoshizawa, M.: Lung Tumor Motion Prediction Based On Multiple Time-Variant Seasonal Autoregressive Model for Tumor Following Radiotherapy, System Integration (SII), IEEE/SICE International Symposium (2010).
13. Ichiji, K., Homma, N., Sakai, M., Takai, Y., Narita, Y., Abe, M., Sugita, N., and Yoshizawa, M.: Respiratory Motion Prediction for Tumor Following Radiotherapy by using Time-variant Seasonal Autoregressive Techniques, 34th Annual International Conference of the IEEE EMBS (2012).
14. Huang, S. H., Tsao, J., Yang, T. C. and Cheng, S.: Model-Based Signal Subspace Channel Tracking for Correlated Underwater Acoustic Communication Channels, IEEE Journal of oceanic engineering, Vol. 39, No. 2 (2014).
15. Man, T., Wong, S. C., Jian, X. and Zhan, G: An Aggregation Approach to Short-Term Traffic Flow Prediction, Intelligent Transportation Systems, IEEE Transactions on, Vol.10, Issue: 1, (2009).
16. Albrecht, D. G. and Geisler W. S.: Motion selectivity and the contrast-responds function of simple cells in the visual cortex, Vis Neurosci, pp. 531 - 546 (1991).
17. Denman, S., Fookes, C. and Sridharan, S.: Improved Simultaneous Computation of Motion Detection and Optical Flow for Object Tracking, Digital Image Computing: Techniques and Applications, pp. 175 - 182 (2009).
18. Sanchez, A. and Rios, H.: Segmentación de objetos en movimiento por flujo óptico y color sin información a priori de la escena, Research in computing science, Avances en Inteligencia Arti_cial, vol. 62, pp. 151 - 160, (2013).
19. Harris, C. and Stephens, M.: A combined corner and edge detector, Proceedings of the 4th Alvey Vision Conference, pp. 147 - 151 (1988)
20. Shi, J. and Tomassi, J.: Good features to track,9th IEEE Conference on Computer Vision and Pattern Recognition (1994).
21. Lucas, B.D. and T. Kanade.: An iterative image registration technique with an application to stereo vision, Proceedings of the 1981 DARPA Imaging Understanding Workshop, pp. 121-130 (1981).
22. Haykin, S.: Nonlinear Method of Spectral Analysis, Springer-Verlag (1979) 23. Kay, S. M., Marple, S.L.: Spectrum Analysis- a modern perspective, Proc. IEEE, vol. 69, pp 1380 - 1419 (1981).
24. Chen, J., Wu, C.: Wu, K.-L., Litva, J.: Combining an autoregressive (AR) model with the FD-TD algorithm for improved computational efficiency, Microwave Symposium Digest, IEEE MTT-S International (1993).

25. Hansun, S.: A New Approach of Moving Average Method in Time Series Analysis, Microwave Symposium Digest, New Media Studies (CoNMedia) (2013)
26. Phillps, P., Perron, P.: Testing fr a unit root in time series regression, biometrika, vol. 7, 2, pp. 335-346 (1988).