Volume 12, Number 3
June 2011 (click for next issue)
Submission Policy, Template and Copyright Form

ISSN 1473-804x Online
ISSN 1473-8031 Print

International Journal of Simulation
Systems, Science & Technology

Guest Editor: Ismail Saad, *Universiti Malaysia Sabah, Kota Kinabalu, Malaysia*
Editor-in-chief: David Al-Dabass**,** *Science and Technology, Nottingham Trent University, UK*

A publication of the United Kingdom Simulation Society

Editorial Policy and Editorial Board | Submission Policy and Author's Instructions

Computational Analysis of Ballistic Saturation Velocity in Low-Dimensional Nano-MOSFET

Ismail Saad[1], Khairul A. M.[1], Nurmin Bolong[1], Abu Bakar A.R[1] and Vijay K. Arora[2]

[1]**N**anoelectron**i**cs Devi**ce** & Mate**r**ial (NiCER) Research Group, School of Engineering & IT, Universiti Malaysia Sabah, 88999, Kota Kinabalu, Sabah
[2]Division of Engineering and Physics, Wilkes University, Wilkes-Barre, PA 18766, U. S. A.
ismail_s@ums.edu.my / ismailsaad07@gmail.com

***Abstract* - The computational analysis of ballistic saturation velocity for low-dimensional nano-devices was presented. The ballistic transport is predicted in the presence of high electric field for non-degenerate and degenerate regime. The saturation velocity is found to be ballistic regardless of the device dimensions. The intrinsic velocity limits this saturation velocity. It's does not sensitively depend on the ballistic or scattering-limited nature of the mobility. In the degenerate realm, the saturation velocity is shown to be the Fermi velocity that is independent of temperature but strongly dependent on carrier concentration. In the non-degenerate realm, the intrinsic velocity is the thermal velocity that depends only on the ambient temperature.**

Keywords - Ballistic transport, Nano-devices, Low-Dimensional devices, Saturation velocity

I. INTRODUCTION

The Metal Oxide Semiconductor Field Effect Transistor (MOSFET) is created when the electric field between the gate and the semiconductor in such that an inverted carrier population is created and forms a conducting channel. This channel extends between the source and drain regions, and the transport through this channel is modulated by the gate potential. As the channel length has gotten smaller, there has been considerable effort to incorporate a variety of new effects into the simple (as well as the more complex) models. These include short-channel effects, narrow width effects, degradation of the mobility due to surface scattering, hot carrier effects, and velocity overshoot. However, as gate lengths have become less than 100 nm, the issue is becoming one of *ballistic* transport rather than these other problems [1- 4]. The speed is determined by the ease with which the carrier (electron or holes) can propagate through the channel of the device. In the earlier designs, the mobility of the carrier was believed to be of paramount importance. However, as development of the devices to nanoscale dimensions continued it became clear that the saturation velocity plays a predominant role [5-7]. The higher mobility brings an electron closer to saturation as a high electric field is encountered, but saturation velocity remaining the same no matter what the mobility. Until today, there is no clear consensus on the interdependence of saturation velocity on low-field mobility that is scattering-limited. There are a number of theories of high-field transport to answer this interdependence. Among them are Monte Carlo simulations, energy-balance theories, path integral methods, green function and many others. Rigor of mathematics and a number of clandestine parameters that are used in these simulations present a foggy picture of what controls the ultimate saturation of drift velocity.

Ballistic transport is referring to the situation in which the channel length is less than the mean-free path of the carriers, so that very little scattering occurs within the channel itself [8-10]. For example if the thermal velocity of a carrier in silicon taken as 2.5×10^7 cm/s at room temperature, a channel mobility of 300 cm^2/V.s leads to a relaxation time of 5×10^{-14} second and a mean-free path of the order of 12×10^{-7} cm, or 12 nm. Thus, only a few scattering events are expected to occurs in a channel length of 20 to 30 nm. While this is a very crude approximation, it points out that the properties of the carriers in these very small devices will be quite different than those in larger devices. In this regards, the paper discussed and elaborated the fundamental theory of ballistic saturation velocity for giving a meaningful interpretation of nanoscale MOSFET.

II. NATORI VS. LUNDSTROM MODEL OF BALLISTIC TRANSPORT

The principle of possibility that MOSFET would operate in ballistic regime as the gate length been scaled towards sub-100nm was purportedly introduced by Natori (1994). However, a simple treatment of ballistic transport as MOSFETs are scaled to their limits and as new devices structure are explored were firstly examined by Lundstrom (1997). Natori developed his expression of I-V characteristics with a full quantum mechanical (QM) basis. The current is expressed in terms of elementary parameters only without depending on the carrier mobility, μ. It is

independent of the channel length L_g and is proportional to the channel width W and oxide capacitance C_{ox}. Due to the quantum effects the energy bands in the channel split and the MOSFET current can be described in terms of energy at each of the split bands using the Fermi Dirac statistics. Thus, the total current can be found by summing the current over all the one-dimensional sub-bands. The sub-channel current component flowing in a direction is given by the product of the unit charge, the number of carriers flowing into the sub-channel per unit time, the transmission coefficient of the sub-channel and the probability that the destination is empty. This sub-channel current component must be integrated over the carrier energy to find the total current. The number of carriers flowing into the sub-channel is further expressed by the product of the input carrier group velocity, the density of states and the probability that the state is occupied by the carrier. The probability of carrier occupancy is given by the Fermi distribution function with the source Fermi level on the source side and that with the drain Fermi level on the drain side of the sub-channel. Both current directions the one from the source to the drain and that in the opposite direction are considered. Thus the total current in the sub-channel is expressed as

$$I = q\sum_{valley}\sum_{n_y}\sum_{n_z}\int\{\vec{\nu}\,\vec{D(E)}\,f(\phi_{FS},E)\left[1-f(\phi_{FD},E\right] - \overleftarrow{\nu}\,\overleftarrow{D(E)}\,f(\phi_{FD},E)\left[1-f(\phi_{FS},E\right]\}T(E)dE \tag{1}$$

where $\vec{\nu}$ is the carrier group velocity of 1D wave propagating towards the drain though a certain sub-channel, $\vec{D(E)}$ is the density of states (DOS) for that wave, both evaluated near the source side edge of the sub-channel. $\overleftarrow{\nu}$ and $\overleftarrow{D(E)}$ are the group velocity and the DOS for carriers propagating from drain to source near the drain side edge of the sub-channel. $f(\phi_F,E)$ is the Fermi distribution function with the Fermi level ϕ_F . The source and drain regions are assumed to be ideal reservoirs with the Fermi energies ϕ_{FS} and ϕ_{FD} respectively. They feed carriers in thermal equilibrium to the channel and also absorb carriers from the channel without reflection (Natori 2006). T(E) is the transmission coefficient of the sub-channel at energy E that represent the probability that a carrier with certain energy can be transmitted over the barrier by the thermionic emission process. In ballistic transport its implies that T(E)=1.

Based on transmission view of the device [11], there is an energy barrier between the source and the drain that prevent the current to flow. This potential energy barrier height is modulated indirectly by the gate voltage V_{GS} which is defined to begin at the top of the barrier. A positive V_{GS} pushes the energy barrier down and allows current to flow. This operational principle works similarly as bipolar transistor which has a similar energy band diagram except that a positive base-emitter voltage V_{BE} lowers the height of the barrier and allows current to increase exponentially [12]. In Lundstrom theory, two major scattering regions can be identified. The scattering in the barrier between the channel and the source, which gives a reflection r_s, and the scattering within the channel, which gives a reflection r_c. In both cases, the reflection coefficients are related to transmission coefficients t by

$$r_s = 1-t_s \quad , \quad r_c = 1-t_c \tag{2}$$

The steady-state flux which reaches the drain a_D can now be written in terms of the injecting flux a_s from source to the source-channel barrier (which is a function of the depth y) as

$$a_D = a_S t_s t_c \tag{3}$$

At the entrance to the channel (which is taken to be $x = 0$, with x the axis aligned from source to drain), the density of carriers can be written as

$$n(0,y) = \left[\frac{t_s a_S + r_c t_s a_S}{\nu_T}\right] = \frac{t_s a_S\left(1+r_c\right)}{\nu_T} \tag{4}$$

where, ν_T is the velocity of the positively and negatively directed fluxes, and y is the direction of the channel depth (normal to the oxide-semiconductor interface). Solving for t_s in these equation yields

$$a_D = n(0,y)\nu_T\left(\frac{t_c}{1+r_c}\right) = n(0,y)\nu_T\left(\frac{1-r_c}{1+r_c}\right) \tag{5}$$

The sheet carrier density is given by integrating over the y coordinate, as

$$n_s = \int_0^{y_{\max}} n(0,y)dy = \frac{C_{ox}}{q}\left(V_{GS}-V_T\right) \tag{6}$$

With this result, the drain saturation current can be written as

$$I_{Dsat} = C_{ox}W\nu_T\left(\frac{1-r_c}{1+r_c}\right)\left(V_{GS}-V_T\right) \tag{7}$$

Equation (7) may be compared with a simple 1D model of a MOSFET. However here the reverse current is represented by the r_c term in the equation, but the form is quite similar to that of the simple theory. Also here the mobility μ is not defined, but instead the carrier transport is discussed in terms of the velocity v_T and the transmission and reflection coefficients within the devices. In the ballistic limit the back scattering coefficient $r_c=0$ and the transmission coefficient $t_c=1$. Thus the maximum current is controlled by the injection velocity at the thermal source.

III. BALLISTIC INTRINSIC VELOCITY

Velocity response to the electric field results in velocity saturation in a high electric field. The current in a resistive (channel) is limited by this saturation value $I_{sat3} = n\ q\ v_{sat}\ A$ (3D), $I_{sat2} = n_s\ q\ v_{sat}\ W$ (2D), $I_{sat1} = n_l\ q\ v_{sat}$ (1D) which in turn depends on the doping concentration n_d ($d=1, 2, 3$) and the saturation velocity v_{sat}. It is, therefore, essential to assess the magnitude of this saturation velocity that results in the current saturation.

In the archived work of Arora (1985) [8], the ballistic (B) transport, although not specifically mentioned by that name, was predicted in the presence of a high electric field. The theory developed was for non-degenerate bulk semiconductors that gave saturation velocity comparable to the thermal velocity. Recently, the theory has been extended to embrace all dimensions under both degenerate and non-degenerate conditions [9]. In equilibrium, the band diagram is flat and randomly oriented velocity vectors cancel each other. As the applied electric field tilts the band diagram, an electron traveling in the direction of electric field finds it difficult to surmount the barrier. An electron traveling in the opposite direction accelerates in a mean free path and collides, randomizing its velocity and restarting its journey for another mean free path.

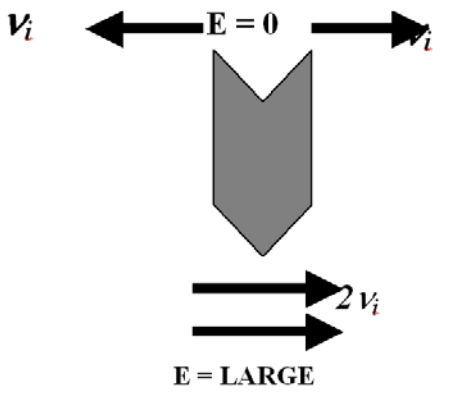


Figure 1 Random intrinsic velocity vectors $\vec{v}_i$ in a zero-field transforming to streamlined vectors in a very high electric field.

The net result is that the random vectors $\vec{v}_i$ streamline in a very high electric field. In the presence of a very high electric field, all electrons are streamlined opposite to the direction of the applied electric field. The ballistic (B) nature of the velocity is apparent from fig. 1 that streamlines the randomly oriented velocity vectors in a very high electric field.

In the absence of external stimulation, the carrier velocity vectors are randomly oriented. This means that their dipole energy $\pm qE_\ell \ell_o$ in a mean free path ℓ_o is also zero. However, as an electric field is applied, the electron energy decreases by $qE_\ell \ell_o$ for electrons drift opposite to the electric field and increases for those drifting parallel to the electric field. This creates two quasi Fermi levels that so-called electrochemical potentials $E_F \pm qE_\ell \ell_o$ within one mean free path on each side of the location of a drifting electron (or a hole). E_F is the Fermi energy related to carrier concentration.

As conduction and valence bands and associated Fermi and intrinsic levels tilt in an electric field, electrons in antiparallel directions are favored over those in the parallel direction finding it difficult to surmount the potential barrier with the net result that all carriers are drifting with the ballistic velocity in an infinite electric field as shown in fig. 2.

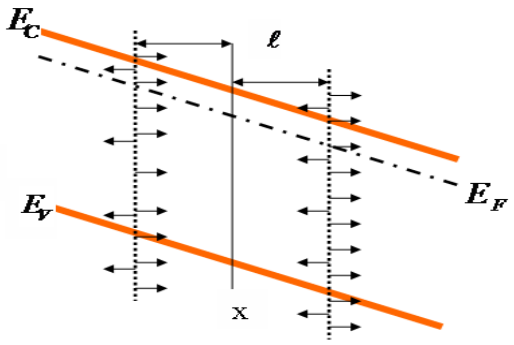


Figure 2 Partial streamlining of random motion of the drifting electrons on a tilted energy band diagram in an electric-field.

From traveling quantum wave perspective, the waves are reflected by the apparently insurmountable barrier in an intense driving field. In this scenario, the drifting electrons are in a relay race, passing on their drift velocity to next electron at each collision, hence the velocity of the electrons is unaffected by the collision and thus called as Ballistic (B) transport of a carrier. This is in direct contrast to the work of those researchers who believe in enhanced scattering for velocity to saturate. Published literature is inconclusive in predicting the dependence of high-field saturation velocity to low-field mobility. Often, the higher saturation velocity is indicated to arise from higher low-field mobility. This

confuses the issue as to what scattering controls the mobility and what controls the saturation velocity. Arora (1985) [8] framework includes all complicated scattering interactions into a single effective mean free path ℓ_o scalable under the strength of carrier scattering that leaves the saturation velocity unaltered and hence B. The mean path may be affected by the onset of quantum emission. The saturation velocity on the other hand is scaled by band structure parameters, unaffected either by the electric field or by the scattering that controls low-field mobility.

The ultimate unidirectional drift velocity is the saturation velocity that is the average of its absolute value $|v| = \sqrt{2E_k/m^*}$, where E_k is the kinetic energy for a given dimensionality and m^* is the carrier effective mass. When this averaging is taken by including the Fermi-Dirac distribution and density of states as a weight for a given dimensionality, the intrinsic B velocity for a semiconductor is obtained as

$$v_{id} = v_{thd} \frac{\Im_{\frac{d-1}{2}}(\eta_{Fd})}{\Im_{\frac{d-2}{2}}(\eta_{Fd})} \tag{8}$$

with

$$v_{thd} = v_{th} \frac{\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)} \tag{9}$$

$$v_{th} = \sqrt{\frac{2k_BT}{m^*}} \tag{10}$$

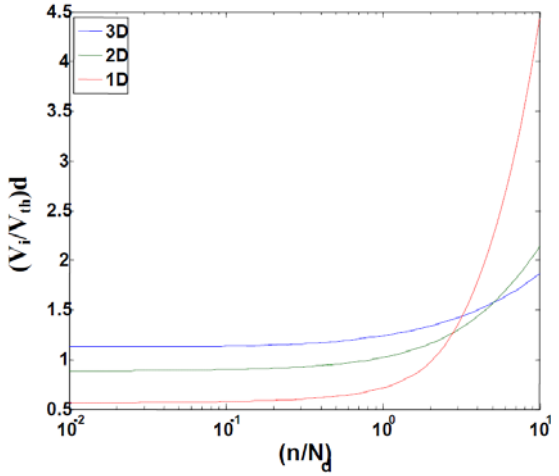


Figure **3** Relative intrinsic velocity versus normalized carrier density.

Figure 3 gives a plot of intrinsic B velocity normalized to the thermal velocity of equation (10). In the non-degenerate regime, this ratio is $2/\sqrt{\pi} = 1.128$ for a 3D semiconductor, it drops to $\sqrt{\pi}/2 = 0.886$ for a 2D nanostructure and further drops to $1/\sqrt{\pi}$ =0.564 for a 1D nanostructure.

IV. BALLISTIC VELOCITY IN NON-DEGENERATE & DEGENERATE REGIME

In the non-degenerate and degenerate regime, the B intrinsic velocity is found to be dependence on the temperature and the carrier concentrations. For the non-degenerate regime, the Fermi factor becomes one since both the Fermi-Dirac integral of order $\Im_{1/2}$, $\Im_0$ and $\Im_{-1/2}$ is following the Maxwell Boltzmann approximation. The B intrinsic velocity for non-degenerate regime can be derived from equation (8) to (10) for each system dimensionality and respectively given as

$$v_{i3} = \sqrt{\frac{8k_BT}{\pi m^*}} \tag{11}$$

$$v_{i2} = \sqrt{\frac{\pi k_BT}{2m^*}} \tag{12}$$

$$v_{i1} = \sqrt{\frac{2k_BT}{\pi m^*}} \tag{13}$$

The normalization of non-degenerate B intrinsic velocity to thermal velocity of equation (10) for d = 3, 2, 1 is as depicted in figure 2.3. It is found that the ratio of thermal velocity is decreasing from 3D to 1- dimensional nanostructure system and is a factor of temperature. The limiting intrinsic velocity for three-dimensional (bulk) non-degenerate electrons is shown to be $2\sqrt{2k_BT/\pi m^*}$ as shown in equation (4) after being simplified. The ballistic velocity reported by Lundstrom and Guo (2005) [10] is half of this value for the reason that the authors are considering only the average over one-half of the Maxwellian distribution. Another difference is that the ballistic velocity $\sqrt{2k_BT/\pi m^*}$ quoted by Lundstrom and Ren (2002) [12] actually applies to bulk case not to the Q2D gas being considered in this study.

For degenerate case, the weighted average $|v|$ of the magnitude of the velocity with Fermi Dirac distribution is

also given by equation (1). However, in the degenerately doped regime, the B intrinsic velocity depends on Fermi velocity v_F instead of the thermal velocity v_{th} given as

$$v_F = \sqrt{\frac{2(E_F - E_C)}{m^*}} \tag{14}$$

The Fermi Dirac distribution part can be simplified for the d = 3, 2 and 1 system as shown by

$$\frac{\Im_{\left(\frac{d-1}{2}\right)}(\eta_d)}{\Im_{\left(\frac{d-2}{2}\right)}(\eta_d)} = \frac{d}{d+1} v_{Fd} \tag{15}$$

Thus, the generic B intrinsic velocity can finally be expressed as

$$v_{id} = \frac{d}{d+1} v_{Fd} \tag{16}$$

The Fermi velocity is derived by using carrier concentrations equation in each dimensionality. Then, the resulted B intrinsic velocity for degenerately doped regime for d = 3, 2, 1 of a low-dimensional system is given respectively as

$$v_{i3} = \frac{3}{4}\frac{\hbar}{m^*}\left(3n_3\pi^2\right)^{1/3} \tag{17}$$

$$v_{i2} = \frac{2}{3}\frac{\hbar}{m^*}\sqrt{2\pi n_2} \tag{18}$$

$$v_{i1} = \frac{1}{4}\frac{\hbar}{m^*} n_1 \pi \tag{19}$$

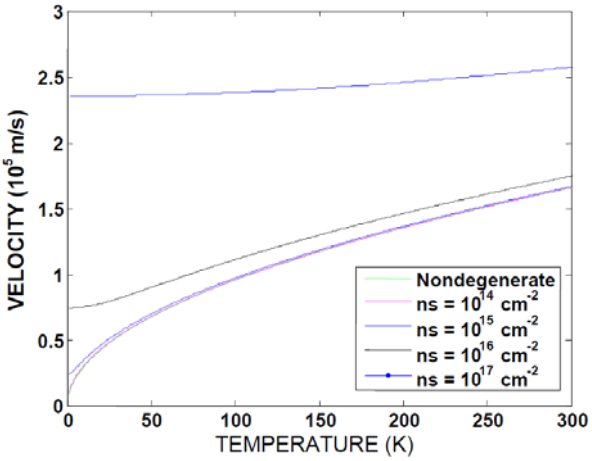


Figure 4. Velocity as a function of temperature for carrier concentration 10^{14} cm^{-2}, 10^{15} cm^{-2}, 10^{16} cm^{-2}, 10^{17} cm^{-2}. Also shown is the non-degenerate limit applicable to low-carrier concentration.

Figure 4 indicates the B intrinsic velocity as a function of temperature. The velocity for low carrier concentration follows $T^{1/2}$ behavior independent of carrier concentration. However for high concentration (degenerate carriers) the velocity depends strongly on concentration and becomes independent of the temperature. Meanwhile, fig. 5 shows the graph of B intrinsic velocity as a function of carrier concentration for three temperatures (T= 4.2 K, 77 K, and 300 K). As expected, at a low temperature, carriers follow the degenerate statistics and hence their velocity is limited by an appropriate average of the Fermi velocity that is a function of carrier concentration.

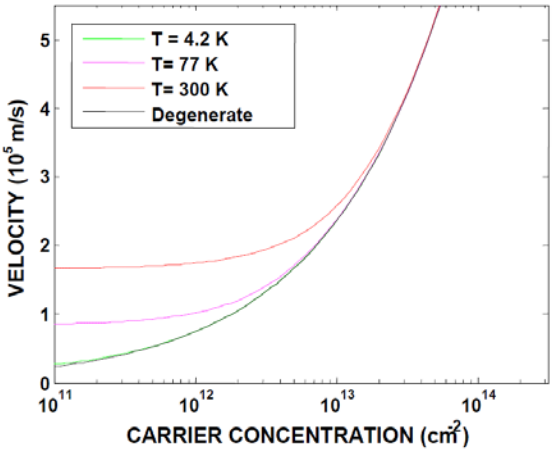


Figure 5 Velocity versus concentration for temperature T= 300, 77, and 4.2 K.

V. CONCLUSION

Using the distribution function that takes into account the asymmetrical distribution of drifting electrons in an electric field is presented. This distribution function

transforms the random motion of electrons into a streamlined one that gives the ultimate saturation velocity that is a function of temperature in nondegenerate regime and a function of carrier concentration in the degenerate regime. The ultimate drift velocity is found to be appropriate thermal velocity for a given dimensionality for no degenerately doped samples. However, the ultimate drift velocity is the appropriate average of the Fermi velocity for degenerately doped samples.

ACKNOWLEDGEMENT

The authors would like to acknowledge the financial support from FRGS (FRG0248-TK-2/2010) and ERGS funds (ERGS0002-TK-1/2011) of Minister of Higher Education Malaysia (MOHE). The author is thankful to the Universiti Teknologi Malaysia (UTM) and Universiti Malaysia Sabah (UMS) for providing excellent research environment in which to complete this work.

REFERENCES

[1] Natori K. "Ballistic metal-oxide-semiconductor field effect transistor". Journal of Applied Physics, 76 (8), pg 4879 - 4890, 1994

[2] Natori K. "Ballistic MOSFET Reproduces Current-Voltage Characteristics of an Experimental Device". IEEE Trans. On Electron Devices, Vol. 23, No. 11, pg 655-657, 2002

[3] Lundstrom M. "Elementary Scattering Theory of the Si MOSFET". IEEE Electron Device Letters, Vol. 18, No. 7, pages 361-363, 1997

[4] Rahman A, Guo J, Datta S and Lundstrom M. "Theory of Ballistic Nanotransistors". IEEE Trans. Electron. Devices, Vol. 50, pp. 1853-1864, 2003

[5] Mugnaini G and Iannaccone G. "Physics-Based Compact model of Nanoscale MOSFETs – Part I: Transition from Drift-Diffusion to Ballistic Transport". IEEE Trans. Electron. Devices, Vol. 52, No. 8, pg. 1795-1801, 2005

[6] Mugnaini G and Iannaccone G. "Physics-Based Compact model of Nanoscale MOSFETs – Part II: Effects of Degeneracy on Transport". IEEE Trans. Electron. Devices, Vol. 52, No. 8, pg. 1795-1801, 2005

[7] Lundstrom M. "On the Mobility versus Drain Current Relation for a Nanoscale MOSFET". IEEE Electron Device Letters, Vol. 22, No. 6, pg 293-295, 2001

[8] Arora V. K,.(1985). *High-Field Distribution and Mobility in Semiconductors*. Japanese Journal of Applied Physics 24, pp 537-545.

[9] Arora V. K, Tan M. L. P, Saad I, and Ismail R. (2007). *Ballistic quantum transport in a nanoscale metal-oxide-semiconductor field effect transistor*. Applied Physics Letter, Vol. 91

[10] Lundstrom M and Guo J. (2005). *Nanoscale Transistor-Device Physics, Modeling and Simulation*,, Springer, 1st edition, New York.

[11] Datta S, Assad F and Lundstrom M. (1998). *The Si MOSFET from a Transmission viewpoint.* Superlatt. Microstruct, Vol. 23, pages 771-780.

[12] Lundstrom. M and Ren. Z. (2002). *Essential Physics of Carrier Transport in Nanoscale MOSFETs.* IEEE Trans. Electron Dev., 49, pp. 133-141.

Design and Simulation Analysis of Vertical Double-Gate MOSFET (VDGM) Structure for Nano-device Application

Ismail Saad, Nurmin Bolong, P. Divya, Bablu K. Ghosh and Kenneth Teo Tze Kin

Nanoelectron**i**cs Devi**ce** & Mate**r**ial (NiCER) Research Group
School of Engineering and Information Technology, Universiti Malaysia Sabah, 88999, Sabah, Malaysia
ismail_s@ums.edu.my/ ismailsaad07@gmail.com

***Abstract* — Design and simulation analysis of vertical MOSFET structure with double gate structure on each side of insulating pillar for nanodevice applications is presented. The body doping effect on vertical channel for channel length, L_g = 50nm and analyzing its effect towards such small devices was successfully performed. The analysis continued with the comparative investigation of device performance with conventional planar MOSFET as scaling L_g down to 50nm. The final part evaluates the innovative design of incorporating dielectric pocket (DP) on top of vertical MOSFET turret with comprehensive device performance analysis as compared to standard vertical MOSFET in nanoscale realm. An optimized body doping for enhanced performance of vertical MOSFET was revealed. The vicinity of DP near the drain end is found to reduce the charge sharing effects between source and drain that gives better gate control of the depletion region for short channel effect (SCE) suppression in nanodevice structure.**

Keywords -Vertical MOSFET, Doping effect, Planar MOSFET, Dielectric Pockets, Short channel effect

I. INTRODUCTION

Scaling the CMOS technology into nanometer regime requires innovative approach in overcoming a number of short channel effect (SCE). Vertical MOSFETs built on the sidewalls of silicon pillars are increasingly being studied as an alternative to standard planar MOSFETs for the scaling of CMOS into the nanometer regime [1 – 3]. With the advantages of controlled gate length by relax photolithographic process, high drive current per unit silicon area and decoupled channel length from packing density, vertical MOSFET in double-gate and/or surround gate structure has become prominent candidate to extend CMOS technology to and beyond the 45nm as depicted by International Technology Roadmap for Semiconductor (ITRS) [4]. The body doping effects, comparative performance with planar MOSFET and innovative design of vertical MOSFET structure are the critical parameters needs to be considered. The vertical MOSFET structure usually utilizes an undoped body for the following reasons. Undoped vertical MOSFETs can avoid the dopant fluctuation effect, which contributes to the variation of the threshold voltage and drive current [5]. The undoped body in vertical MOSFETs can enhance the carrier mobility due to the absence of depletion charges which can significantly contribute to the effective electric field, thus degrading the mobility [6]. However, without body doping as a tool to adjust the threshold voltage, undoped vertical MOSFETs need to rely on gate work function to achieve multiple threshold voltages on a chip. Tunable metal gate technology thus needs to be developed for double gate (DG) vertical MOSFETs. However, a metal gate with a tunable work function has not been integrated in DG MOSFETs due to technological difficulties [7 – 8]. Therefore, body doping remains as an alternative to set appropriate threshold voltages for DG vertical MOSFETs. The incorporation of so-called dielectric pocket (DP) for limiting dopant diffusion and controlling SCE such as threshold voltage (V_T) roll-off have been proposed and fabricated for planar [9] and vertical MOSFET [10 - 11]. However the epitaxial process [12] requires critical lithography and is not CMOS compatible. This paper presents the innovative design consideration of vertical MOSFET architecture that was successfully being evaluated in the context of body doping effect, performance evaluation with conventional planar MOSFET and the effects of dielectric pocket (DP) structure as scaling the channel length into nanoscale realm.

II. VERTICAL MOSFET TECHNOLOGY

In order to extend CMOS scaling to its physical limits while maintaining performance improvements, devices with structure different from the conventional single gate MOSFET will have to be introduced. A most promising concept in this direction is the double gate MOSFET. This structure has several advantages over its single gate counterpart:

- If perfect coupling between front and back gate is achieved, short channel effects can be suppressed and devices featuring an ideal sub-threshold swing of 60mV/decade can be fabricated
- The body doping can be reduced dramatically, thus improving carrier mobility
- The channel width per unit area is at least doubled, thus increasing the current drive per unit area of the transistors.

Double gate MOSFETs can be grouped in three classes:

- **Planar Double Gate and Gate All Around (GAA) MOSFETs**. In planar double gate devices both the current-carrying plane and the current flow are parallel to the wafer surface. In GAA MOSFETs the current can flow on planes perpendicular to the wafer surface as well.
- **FinFETs.** In these devices the current-carrying plane is perpendicular and the current flow parallel to the wafer surface.
- **Double and Surround Gate Vertical MOSFETs**. In these devices both the current-carrying planes and the current flow are perpendicular to the wafer surface.

The most notable feature of vertical MOSFETs is that their channel length is defined by non-lithographic methods, such as ion implantation or epitaxial deposition. As a result, shorter channel length is possible in nanometer regime. On top of that, the gate length is decoupled from the packing density. This feature is very useful for low-leakage and DRAM applications. Moreover, the fabrication of double and surround gate MOSFETs is straightforward in vertical technology, so that the channel width per unit area is dramatically increased in comparison with planar devices. Consequently, this would increase the drive current per unit area. Finally, vertical MOSFETs with sub-100nm pillar or ridge width suppress short-channel effects and drain leakage.

The drawbacks of vertical MOSFETs are due to their low compatibility with planar CMOS technology, particularly when epitaxial method is employed. A precise control of the channel length is generally difficult by ion implantation though much easier by epitaxy. Another issue is the gate/drain and gate/source parasitic capacitance as the top drain electrode generally is not self-aligned to the gate. In addition, for a thin channel fully depleted (FD) devices other processing challenges arise for short channel effects (SCE) control. High resolution lithography is needed for pillar or ridge width definition.

III. DEVICE STRUCTURE AND MODELS

The simulated vertical MOSFET structure [3] is shown in Fig.1 with the double gate region (in contact), drain and source electrode, channel length (L_g), oxide thickness (t_{ox}), body or channel and the respective dimensions of the device are explicitly shown. Notice that an electrode line is visible in Fig.1 to make sure that the left and right gates are in contact. A uniform doping profile is assumed and applied to drain (n-type), source (n-type), double gate (n-type) and body (p-type) of the device with the concentration of $1x10^{20}$ cm^{-3}, $1x10^{20}$ cm^{-3}, $1x10^{21}$ cm^{-3} and 3.5×10^{18} cm^{-3} respectively. The DP (not shown) is placed near the drain-end contact.

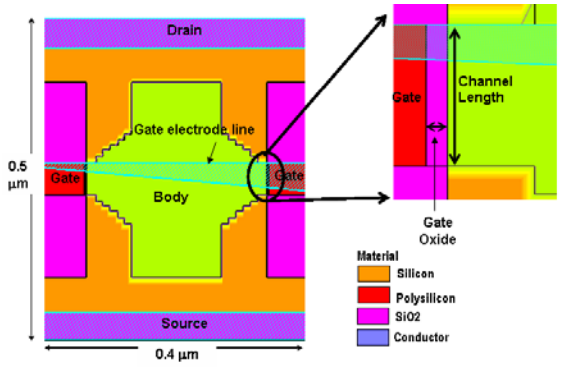


Figure1 Vertical MOSFET structure shows double gate, source, drain, body, channel length and gate oxide

For numerically computing the electrical behavior of the device, an appropriate physics based models has to be evoked carefully. There are five categories of models involved namely mobility, carrier generation and recombination, carrier statistics, impact ionization and tunneling models. The inversion layer mobility model [13] was employed for its dependency on the transverse field (i.e field in the direction perpendicular $E_{\perp}$ to the Si/SiO_2 interface of the MOSFET) and through velocity saturation at high longitudinal field (i.e field in the direction from source-to drain parallel E to the Si/SiO_2 interface) combined with SRH (Shockley-Read-Hall Recombination) with fixed carrier lifetimes models. This recombination model was selected since its take into account the phonon transitions effect due to the presence of a trap (or defect) within the forbidden gap of the semiconductor. An interface fixed oxide charge of $3x10^{10}$ C is assumed with the used of n-type polysilicon gate contact for the device. The Drift-Diffusion transport model with simplified Boltzmann carrier statistics is employed for numerical computation of the device under study.

IV. BODY DOPING EFFECTS ANALYSIS

For body doping effects analysis, three variant of body doping are used: low doped ($N_A = 1x10^{18}$ cm^{-3}), moderately doped ($N_A = 2x10^{18}$ cm^{-3}) and the high doped ($N_A = 3.5x10^{18}$ cm^{-3}). Due to a decrease in body doping, the V_T value also decreases from 0.56V for high doped to 0.36V in moderate doped and to a lower value of 0.15V in low doped body. However, a decreased in doping will ultimately increased the leakage current from 2pA/µm to 7nA/µm and finally to a value of 80µA/µm as shown in fig. 2. Nevertheless, the increased in drive current is almost unity with a value of 7µA/µm to 10µA/µm and 1mA/µm respectively. These effects arise due to the fact that at higher doping the surface mobility is decreased and a better gate electrostatic potential observed within the device which makes the leakage current controllable. However, as the doping level decreased, the carrier mobility is increased and consequently the leakage current will also rise sharply.

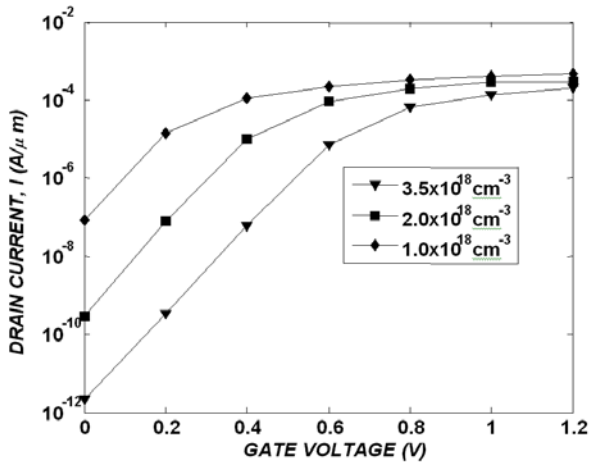


Figure 2 Subthreshold characteristic showing an increased in leakage and drive current with a lower doping level and an increase in threshold voltage with higher doping level.

Due to double gate structure, the increased in drain saturation current (I_{Dsat}) was observed as shown in fig. 3. Even though the I_{Dsat} is high with lower doping level, a high I_{OFF} = 80μA/μm is extremely unacceptable. These results are in conjunction with the value of sub-threshold voltage, which is 89 mV/decade for higher doped, 83 mV/decade in moderate doped and sharply increased to 110 mV/ decade in lower doped device. Thus, an optimize value of body doping is highly vital in order to have a high drive current while maintaining the acceptable leakage current and controlling the aggravated SCE [14]. Failure to control such parameters, the transistor designed will not succeed to work in a giga-scaled integrated circuit where the total standby power of the system is of paramount important.

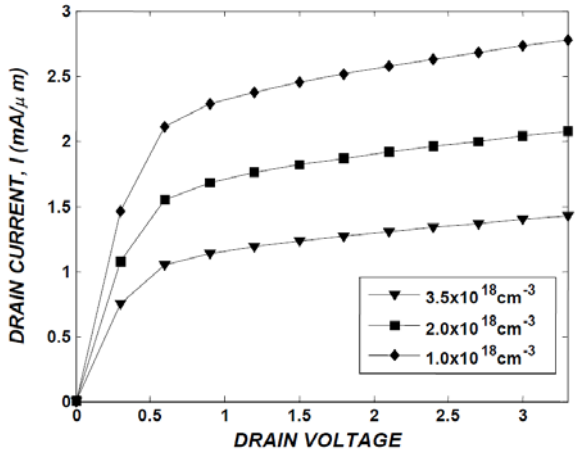


Figure 3 Output characteristic shows an increased in drain saturation current with a lower doping level in a double-gate configuration.

V. COMPARATIVE ANALYSIS

Compatible process parameters for planar and vertical MOSFET are maintained such that a valid comparison is made. The channel and source/drain doping is $5x10^{18}$ cm^{-3} and $1x10^{20}$ cm^{-3} respectively with junction depth of between 100nm to 120nm. The conductance of insulating film is considered negligible by using a t_{ox}=5nm and silicon body thickness t_{si}=136nm for both devices.

For scaling the channel length of vertical MOSFET, the height of silicon pillar during dry etch process is vary from 200nm to 300nm. Such heights with fixed nitride thickness of 97nm will approximately prepared 100nm to 50nm channel length. In contrast, for scaling the planar MOSFET L_g, the polysilicon gate length needs to be scaled accordingly before the source/drain ion implantation took place. This process is highly dependent on the accuracy of etching the polysilion gate length by the critical lithography process and limits by the wavelength of the light. In vertical MOSFET with 0.22μm silicon pillar height, a channel length of 50nm is reachable presume relax lithography dependent. The 0.22μm is easily can be done using normal photolithography steps without the needs of using the expensive electron beam lithography. Figure 4 shows the V_T versus the channel length for both devices. As the L_g is decreased, the V_T decreases due to short channel effect (SCE). However, sharply decrease of V_T roll-off is seen to happen for planar device as compared to vertical device. This is due to the double gate structure on both side of vertical channels that makes a better electrostatic control of the channel by the gate even when the channel is scaling down to 50nm.

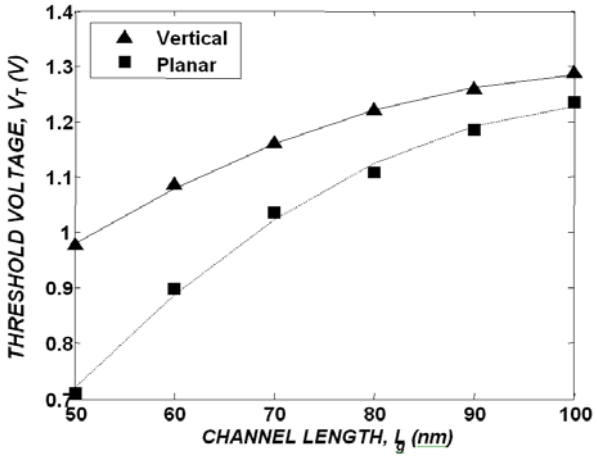


Figure 4 Threshold voltages, V_T roll-off characteristics for planar and vertical MOSFET down to 50nm channel length, L_g

The leakage current (I_{OFF}) defined as the drain current at V_{GS}=0V and low drain voltage (V_{DS}) was also extracted for both devices. This current which is due to the reverse-biased p-n junction at the drain region is a very important parameter in making sure the immobile state of the transistor in sustaining the power drainage. It's particularly essential as the number of transistors per chip growth monotonically in support of system-on-chip (SOC) paradigm.

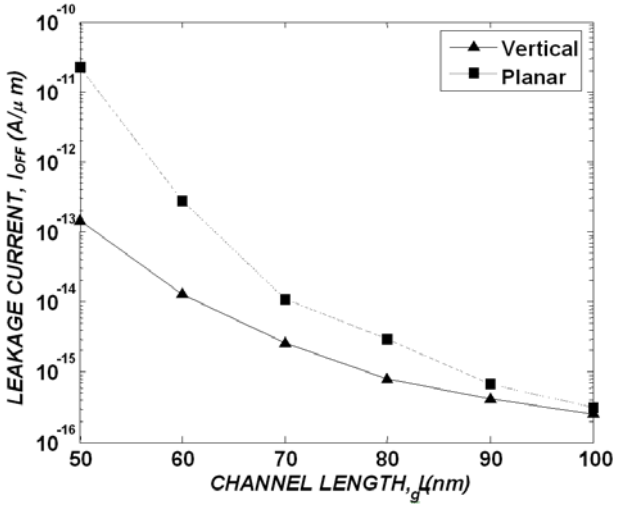


Figure 5 Drain leakage current I_{OFF} for both planar and vertical MOSFET extracted at V_{GS}=0V and low V_{DS}

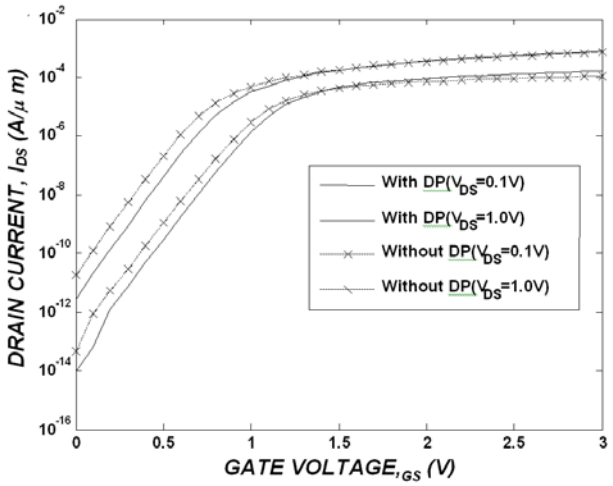


Figure 6 Sub-threshold characteristics of vertical MOSFET with dielectric pocket (DP) and without DP for L_g = 50nm, t_{ox} = 5nm, N_A = 1x10^{18} cm^{-3} at drain voltage V_{DS} = 0.1V and 1.0V respectively.

Figure 5 illustrate the comparison of these values for both devices when the channel is scaling down to 50nm. With the same body doping, the SCE is pronouncedly observed when the channel is scaled to 50nm for both devices. However, due to the structure of vertical MOSFET during the drain-on-top (DOT) mode the leakage current is lower compared to planar device. The I_{OFF} increase when Lg scaled to 50nm and planar MOSFETs have risen up to two decade higher than vertical MOSFETs. For 50nm, the I_{OFF} = 2.3x10-11 A/μm and 1.47x10-13 A/μm is observed for planar and vertical MOSFETs respectively. This gives an advantage to vertical channel device for making sure a lower value of I_{OFF} in an application such as DRAM and SRAM circuit of a planar device.

VI. DIELECTRIC POCKET (DP) ANALYSIS

The Dielectric Pocket (DP) provides a number of purposes. It greatly reduces the influence of large area parasitic bipolar transistor (PBT) in the vertical structure. Reducing the electrical bulk punchtrough effects by preventing the encroachment of the doping from the extrinsic drain and reduces the charge sharing effects associated with the reverse-biased drain to improve threshold voltage control [10]. Figure 6 shows the sub-threshold or transfer characteristics for vertical MOSFET with and without DP for L_g = 50nm, t_{ox} = 5nm, N_A = 1x10^{18} cm^{-3}. Both devices shows an excellent on-off characteristics, however the off-state leakage I_{OFF} in the DP device is lower than without DP device in the drain voltage V_{DS}=0.1V and 1.0V. The I_{OFF} at V_{GS}=0V and V_{DS}=0.1V for DP device is 9.63 x 10^{-15} A/μm and for without DP the I_{OFF} = 4.57 x 10^{-14} A/μm. This is possibly due to the decreases of larger electric field at the drain end reverse-biased PN junction in the vicinity of dielectric pocket.

The decrease of electric field will reduce the leakage current predominately. The leakage current was originated from the body-drain depletion region due to the reverse-biased body-drain PN junction [15]. Similarly the DIBL effect is also observed to be lower in DP device. In DP device the DIBL is 144 mV/V whereas for without DP the DIBL increased to 177 mV/V. These results indicate that DP structure is essentially needed for the suppression of SCE.

The output characteristics of vertical MOSFET with and without DP is shown in fig. 7 for L_g=100nm. The drive current in DP device was shown to be effectively higher than without DP device. Due to the presence of DP in the vicinity of drain region the larger electric field at reverse-biased PN junction is reduced effectively. The reduced electric field E at drain end increased the carrier mobility μ in the channels and ultimately improved the drift velocity v as given by

$$v = \mu E \tag{1}$$

$$\mu = \frac{q\tau}{m^*} \tag{2}$$

where q is the charge, τ is the mean free time between collisions and m^* is the carrier effective mass. Thus the increase in carrier velocity makes the saturation velocity to be higher in DP compared to without DP device.

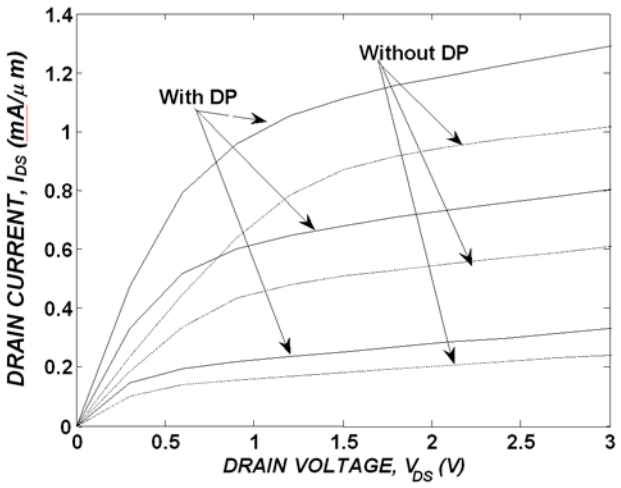


Figure 7 The output characteristics of vertical MOSFET with DP and without DP for channel length L_g = 100nm at V_{GS} = 2.2V, 3.3V and 4.4V respectively.

VII. CONCLUSIONS

The body doping effects analysis of vertical replacement gate (VRG) MOSFET structure found that higher body doping is essential needed for controlling SCE. However, it will decrease the carrier mobility and leads to lower value of drive current. In contrast, low value of body doping will eventually increase the drive current and leads to an elevated increase of leakage current. Thus, an optimize value of body doping is highly vital in order to have a high drive current while maintaining the acceptable leakage current and controlling the aggravated SCE. With moderate body doping of N_A = 2.0x10^{18} cm^{-3}, an acceptable V_T = 0.36V, I_{OFF} = 7nA/μm, SubV_T = 83 mV/decade and I_{ON} = 10μA/μm was successfully obtained for the simulated VRG device structure. The performance comparison with planar MOSFET has reflected the benefit of double-gate (DG) vertical MOSFET. The DG structure over the side of silicon pillar has been shown to have a very good electrostatic gate control over the channel, enabling gate length scaling down to 50nm for vertical MOSFET. The analysis shown that the vertical MOSFET have two decade lower of leakage current, a factor of three larger of drain saturation current and almost ideal value of sub-threshold swing. The innovative design of incorporated dielectric pocket (DP) over the silicon pillar has advantages of detaining higher reverse-biased PN junction electric field for minimizing the leakage current and reducing the charge sharing effects between source and drain that gives better gate control of the depletion region.

ACKNOWLEDGEMENT

The authors would like to acknowledge the financial support from FRGS (FRG0248-TK-2/2010) and ERGS funds (ERGS0002-TK-1/2011) of Minister of Higher Education Malaysia (MOHE). The author is thankful to the Universiti Malaysia Sabah (UMS) for providing excellent research environment in which to complete this work.

REFERENCES

[1] Jayanarayanan S. K. (2004). *Silicon-Based Vertical MOSFETs.* PhD Thesis – University of Texas at Austin.

[2] Mori K, Duong A and Richardson W.F. (2002). *Sub-100-nm Vertical MOSFET With Threshold Voltage Adjustment.* IEEE Trans. Electron Devices, 49(1):pg 61–66.

[3] Hergenrother J. M, Oh S.H, Nigam T, Monroe D, F. Klemens P, and Kornblit A. (2002). *The vertical replacement-gate (VRG) MOSFET.* Solid-State Electronics, 46: pg 939–950.

[4] International Roadmap for Semiconductor 2007 (ITRS, 2007) – *Process Integration, device and structure (PIDS).* http://public.itrs.net.

[5] Liu H, Sin J. K. O, Xuan P and Bokor J. (2004). *Characterization of the Ultra-Thin Vertical Channel CMOS Technology*. IEEE Trans. Electron Devices, 51(1):106–112.

[6] Gili E, Kunz V.D, de Groot C.H, Uchino T, Ashburn P, Donaghy D.C, Hall S, Wang Y, Hemment P.L.F (2004). *Single, double and surround gate vertical MOSFETs with reduced parasitic capacitance.* Solid-State Electronics, Vol 48: pg 511–519.

[7] Schaeffer J K, Capasso C, Fonseca L R C, Samavedam S, Gilmer D C and Liang Y. (2004). *Challenges for the integration of metal gate electrodes.* Int. Electron Devices Meeting (IEDM) pp 287–290

[8] Jeon I. S et al (2004). A novel methodology on tuning work function of metal gate using stacking bi-metal layers. Int. Electron Devices Meeting (IEDM) pp 303–306

[9] Jurczak M, Skotnicki T, Gwoziecki R, and *et al*,. (2001). *Dielectric pockets – a new concept of junctions for Deca-Nanometric CMOS devices*. IEEE Trans. Electron Devices, Vol. 48(8), pg 1770-1774

[10] Donaghy D, Hall S, de Groot C.H, Kunz V.D and Ashburn P. (2004). *Design of 50-nm Vertical MOSFET Incorporating a Dielectric Pocket*. IEEE Trans. Electron Devices, Vol. 51 (1).

[11] Jayanarayanan S. K, Dey S, Donnelly J.P and Banerjee S.K. (2006). *A Novel 50nm Vertical MOSFET with a dielectric pocket.* Solid-State Electronics, Elsevier, pages 897-900.

[12] Jayanarayanan S. K. (2004). *Silicon-Based Vertical MOSFETs.* PhD Thesis – University of Texas at Austin.

[13] Lombardi C. et al., (1988). A Physically Based Mobility Model for Numerical Simulation of Nonplanar Devices. IEEE Transactions on Computer-Aided Design, vol.7, no.11

[14] Lu H, Lu W.Y and Taur Y. (2008). *Effect of body doping on double-gate MOSFET characteristics*, Semicond. Sci. Technol. Vol. 23

[15] Gili E, Kunz V.D, Uchino T, Al Hakim M. M, de Groot C.H, Ashburn P and Hall S. (2006). *Asymmetric Gate-Induced Drain Leakage and Body Leakage in Vertical MOSFETs with Reduced Parasitic Capacitance.* IEEE Transactions on Electron Devices, Vol. 53(5).

Charge Property Modeling of Nanofiltration Hollow Fiber Membranes

Nurmin Bolong*[a,b,c], Ismail Saad [a], A. F. Ismail.[b], M. R. Salim [c], D. Rana.[d] and T. Matsuura [d]

[a] **N**anoelectron**i**cs Devi**ce** & Mate**r**ial (NiCER) Research Group, School of Engineering & IT, Universiti Malaysia Sabah (UMS), Kota Kinabalu, Malaysia
[b]Advanced Membrane Technology Research Centre (AMTeC), Faculty Of Chemical & Natural Resources Engineering, Universiti Teknologi Malaysia (UTM), Johor, Malaysia
[c]Dept. of Environmental Engineering, Faculty of Civil Engineering, Universiti Teknologi Malaysia (UTM), Johor, Malaysia
[d]Industrial Membrane Research Laboratory, Dept. of Chemical and Biological Engineering, University of Ottawa, Canada

*Corresponding Author: nurmin@ums.edu.my

Abstract **- The development of models that predict membrane performance has contributed a better understanding of the basic principles and mechanisms of solute rejection and deposition. It also serves as fundamental properties that allow specific characterization determination. This work fabricates hollow fiber membranes using Polyethersulfone (PES). The membranes were fabricated in-house using phase inversion technique by modification with synthesized charged-surface modifying macromolecules (cSMM). The cSMM comprise with end-group component of Hydroxybenzene sulfonate or Hydroxybenzene carboxylate. The electrical properties of the membranes were modeled by utilizing the combination of irreversible thermodynamic model, Steric-Hindrance Pore (SHP) model and Teorell-Meyer-Sievers (TMS) model. The negatively-charged of the modified hollow fiber membranes was calculated based on sodium chloride rejection experimental performance. The analysis of the modeling results revealed that the modification of nanofiltration membrane using sulfonate induce negative 1.61 electrical properties compared to carboxylate that is negative 1.49 for both type modified PES membranes.**

Keywords - mathematical modeling, negative charge, surface modification, hollow fiber membrane, ionic separations.

I. INTRODUCTION

Main membrane separation technology in water and wastewater treatment are categorized into four classes according to the separation process, namely, reverse osmosis (RO), Nanofiltration (NF), ultrafiltration (UF), and microfiltration (MF). Membrane processes becoming more popular because the processes can disinfect water without chemical additions and avoid the formation of toxic disinfection byproducts [1]. Furthermore, membrane has received more interest in recent years due to increasingly stringent standard requirement for water supply and effluent discharge. The benefits of membrane treatment processes has been highlighted as; compact, modular construction, minimum of moving parts with low maintenance requirements, no chemical addition requirements and minimal chemical sludge disposal, absolute barriers to particles and pathogens, constant filtered water quality irrespective of feed water quality, easy system upgrading and suitable to small systems and distributed locations [2,3].

The success of any separation system involving membrane depends on the quality and suitability of the membrane incorporated in the system. Nanofiltration (NF) process is a better option than UF due to its capability to reject matters with molecular weight (MW) greater than approximately 200Da, yet it requires more energy consumption during operation [4,5]. On the other hand, the relatively coarse pore size distribution of UF made it inadequate to remove organic matter with several hundreds of MW [6]. In addition, nanofiltration separation process involves a combination influence of charge interaction and size exclusion [7].

Advancement in membrane fabrication has pushed the border of membrane limitation and improves their performance into various of application, which makes membrane formation and fabrication is an interesting subject of study. The ability of producing own made membrane allows a greater chance to produce and achieve a successful treatment system since during the production process of membrane it can be tailored to suit the requirements. However, many important factors need to be taken into account and properly controlled in order to allow production of membranes that meet the desired characteristics.

Surface membrane properties are one of the key controls used to optimize membrane performance, especially via membrane surface modification. In membrane separation processes, membrane charge correlated with electrostatic interactions between membrane and solutes [8-10]. The negative properties of membrane surface usually are contributed by the sulfonic and/or carboxylic acid groups which are deprotonated at neutral pH [11,12]. Several researches has studied the membrane surface charge dependence to feed water matrix including factor of pH [13,14], electrolyte concentrations [15] and presence of natural organic matter [16]. All was agreed that by increasing negative surface charge of membranes will increases rejection performance, especially for negatively charged compounds [17].

Surface modification of membrane by blending method has an advantage due to easy fabrication and a single step process. However, the success of surface modification by blending is highly correlated to many factors and still a challenging field of studies. The combination of different nanofiltration membrane removal mechanism namely; convection, diffusion (sieving) and charge effects, offers an advantage towards ionic solutes removal. However, in order to predict and optimize charge nanofiltration membrane performance, the charge properties of the fabricated nanofiltration membranes need to be characterized.

The limitation of specialized instrumentation on the measurement of hollow fiber membrane charge properties is overcome by the mathematical model study. The modeling incorporates retention experiments using ionic solutes. The determination of the membrane charge parameters based on the theoretical models are presented in this work. Therefore, the aim of this paper is to evaluate the influences of charged-surface modifying macromolecules modification on the polyethersulfone (PES) hollow fiber membranes for ionic separation and characterized based on the mathematical models.

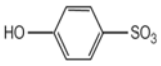
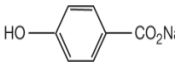
II. EXPERIMENTAL METHODS AND MATHEMATICAL MODEL

A. Hollow fiber membranes

The membranes used in this work were in-house fabricated, using *N*-methyl-2-pyrrolidinone (NMP, >99%) (Merck, Darmstadt, Germany) as solvent. The PES (RADEL A-300, Amoco Chemicals) hollow fiber membrane were modified with charged surface modifying macromolecules (cSMM). The cSMMs are oligomeric glycol polymers synthesized by polyurethane chemistry and tailored with end groups. Hypothetically, the charges were contributed by the end-capped group of cSMM. The end-groups used in this study were sulfonate and carboxlate ion.

The morphology observation of the hollow fiber membranes was made by Field emission scanning electron microscopy (ZEISS SUPRA 35VP). The fiber samples were prepared by immersing the fiber samples in liquid nitrogen and then fractured to obtain a clean break. Table 1 shows the fabrication process and spinning conditions, which were named as HF1 and HF2.

Table 1. Technical fabrication condition of the hollow fiber membranes

Technical fabrication process	Membranes	
	HF 1	HF 2
Dope formulation (wt%)	PES/PEGHBS (22/1)	PES/DEGHBC (22/1)
cSMM end-group formula	Hydroxybenzene sulfonate 	Hydroxybenzene carboxylate 
Spinning conditions:		
Dope flow rate:	3.0 cm^3/min	3.0 cm^3/min
Bore fluid composition:	Distilled water	Distilled water
Dimensions of spinneret (mm):	ID/OD 0.3/0.6	ID/OD 0.3/0.6
Air gap distance	5 cm	5 cm

B. Permeation experiments

The permeation experiments were carried out in a cross-flow filtration setup as described elsewhere [18]. Each of the membranes was first permeated with distilled water to check its permeability. Then for all three membranes, permeation experiments were carried out with 0.01 M NaCl (Merck, Darmstadt, Germany). The ionic solute rejections were analyzed using digital conductivity meter (WTW handheld meter model LF330).

C. Mathematical modeling

The Extended-Nerst Planck Equation (1) forms the basis for the description of the transport of ions or solutes inside the membranes. The equation (1) which includes Donnan expression; describes the partitioning of solutes between solution and membrane [7]. The terms represent transport due to diffusion, electric field gradient and convection respectively.

$$J_{si} = -\left(D_{i,p}\frac{dC_i}{dx}\right) - \left(\frac{z_i S C_i D_{i,p}}{R\cdot T}\cdot F\cdot\frac{d\psi}{dx}\right) + \left(K_{i,c}C_i J_s\right) \quad (1)$$

where J_{si} is the flux of an ion i (mol/m^2.s), $D_{i,p}$ is the diffusivity of the ion i in the membrane (m^2/s), z_i is the valence of the ion i (dimensionless), c is the ion concentration (mol/m^3), G is the gas constant (J/K.mol), F is the Faraday constant (9.648x10^{-4} C/mol), ψ is the electrical potential (V) and $K_{i,c}$ is the convective hindrance factor of membrane (dimensionless).

The solution of the above model requires three structural and electrical parameters namely: pore radius, r_p, effective ratio of membrane thickness to porosity, Ax/Ak, and the effective charge density, X_d. These parameters can be obtained by fitting the rejection data of uncharged and charged solutes. Once these parameters are obtained, the model can be used to predict the separation performance of ions or charged solutes in the system.

Based on Irreversible thermodynamics model which is used to describe the transport phenomena in the process in NF, Spiegler and Kedem [23] further expressed the flux of solute, J_s in a differential form when high concentration difference between retentate (rejection) and permeate exist, which shown in equation (2). By integrating the equation (2) across membrane thickness, equation (3) and (4) were derived. In this study, it is assumed that the real rejection is same as the observed rejection due to the low concentration of solute used. The solute concentrations at the membrane surface, C_m and in the bulk solution, C_b are assumed to be the same. It is because the effect of concentration polarization which decreases the driving force is not obvious at a very low concentration of aqueous solute solution used [24]. With this assumption, the equation was rearranged and described rejection (R) as equation (5).

$$J_s = -P'\left(\frac{dc}{dx}\right) + (1-\sigma)J_v C \tag{2}$$

$$R = 1 - \left(\frac{C_p}{C_m}\right) = \frac{\sigma(1-F)}{1-\sigma F} \tag{3}$$

$$\text{With} \quad F = \exp\left(-\left(\frac{1-\sigma}{P}\right)J_v\right) \tag{4}$$

$$R = \frac{\sigma\left(1-\exp\left(-\frac{1-\sigma}{P}J_v\right)\right)}{1-\sigma\exp\left(-\frac{1-\sigma}{P}J_v\right)} \tag{5}$$

Where P is solute permeability, C_p is solute concentration in the permeate, C_m is solute concentration in the fluid at feed on membrane interface, σ is reflection coefficient, J_v is the flux volume and C is the mean concentration over the membrane thickness.

III. RESULT AND DISCUSSION

The results of filtration performance are shown in Figure 1. The permeate flux versus the applied pressure for all the three membranes obtain a linear profile. The slope of the lines represents the water permeability value for each membrane. As shown, the pure water fluxes of cSMM blended membranes are significantly higher than the unblended membrane at different operating pressures, i.e. the normalized pure water fluxes for HF1 and HF2 were 4.6 and 5.5 L/m^2.h.Bar, respectively. These permeability values are relatively low but still within the range normally obtained with commercial NF membranes available which range from 1.3 to 50.5 L/m^{2}h.Bar [21].

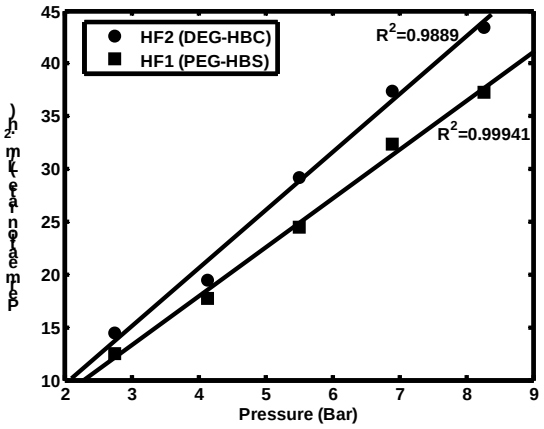


Figure 1. Water permeation performance of cSMM modified membranes at applied pressure range of 2.8-8.3 Bar

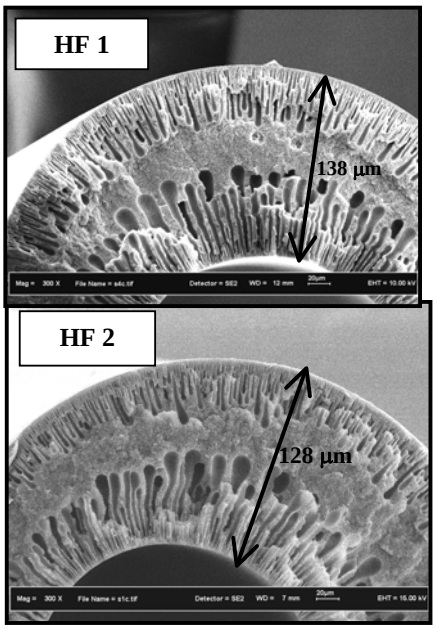


Figure 2. FESEM pictograph of hollow fiber cross-section at 300x magnification

The reason of DEGHBC (HF2) has higher pure water fluxes than PEGHBS (HF1) could be primarily attributed to the physical and morphology of fabricated membranes as observed via FESEM. From Figure 2, the FESEM pictograph estimated that the membrane thickness of HF2 is 127.7 μm whereas for HF1 is 138.3 μm. Hence HF1 have slightly thicker resistance in permeability compared to HF1, since the permeation occurs from outer to inner hollow fiber membrane.

The electrical property of self-fabricated membranes was determined using Teorell-Meyer-Sievers (TMS) model (equation (6)). TMS model is a rigorous approach that has been widely used to describe the membrane electrical properties (in the case of negatively charged membrane) by assuming a uniform radial distribution of fixed charges and mobile species [22]. With the combination with the extended Nernst-Planck model and the Donnan equilibrium theory, the parameters experimentally obtained (σ_{salt}) can be applied to determine the electrical property of membrane using NaCl electrolyte as aqueous feed solution, as follow:

$$\sigma_{salt} = 1 - \left(\frac{2}{(2\alpha - 1)\xi + \sqrt{(\xi^2 + 4)}} \right) \quad (6)$$

Where

$$\alpha = \frac{D_{cation}}{D_{cation} + D_{anion}} \quad (7)$$

With α represent as the transport number of cations in free solution, D as the diffusion coefficient and ξ is the electrical properties of the membrane (negative unit).

σ_{salt} is determined directly from the experimental data of rejection (R) as a function of 1/Jv using one of the best-fitting methods (using MATLAB R14), which is shown in Figure 3. This is derived based on the irreversible thermodynamic model [23,24]. Based on the irreversible thermodynamic model, membrane is treated as a black box as no information about the membrane properties can be obtained from it. Hence, from the derive equation of (5), it is assume that the rejection of NaCl, (R) becomes equal to the reflection coefficient, σ (maximum solute rejection) when the volume flux, J_v is infinite as the filtration flow overtakes solution diffusion. Therefore, σ and P can be determined directly from the experimental data of R as a function of 1/Jv using one of the best-fitting methods.

In Figure 3, the rejection of 0.01M NaCl versus 1/Jv for both membranes was plotted. The electrical properties were then calculated based on the 0.01M NaCl rejection results, which is summarized in Table 2. It should be noted that the higher the electrical property (ξ), indicates the more negative charge and hence better removal of charge solutes. Shown from the results, addition of cSMM using PEGHBS or DEGHBC has increased the electrical properties compared to pure PES (ξ=-0.5) [25]. For both cSMM types, it was found that the modification using cSMM of PEGHBS has higher charge compared to DEGHBC at this particular fabrication condition.

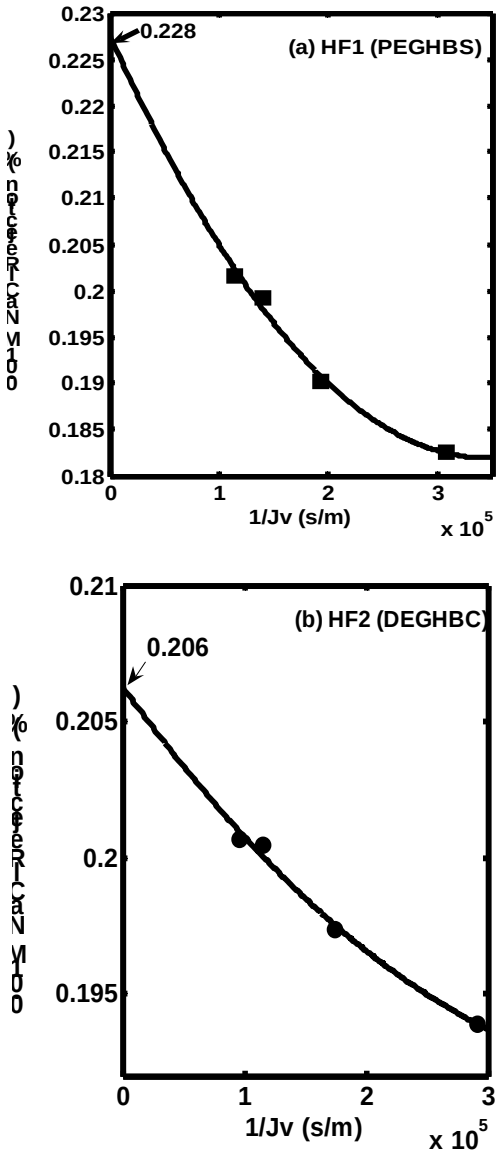


Figure 3. 0.01M NaCl Rejection performance of HF1 and HF2 towards flux

Table 2. Rejection mechanism parameter

Membrane	σ_{salt} (%)	D_{cation} (m^2/s)	D_{anion} (m^2/s)	Charge, ξ (negative)
HF1 (PEGHBS)	0.228	$1.33x10^{-9}$	$2.03x10^{-9}$	1.61
HF2 (DEGHBC)	0.206	$1.33x10^{-9}$	$2.03x10^{-9}$	1.49

IV. CONCLUSION

The first attempted PES hollow fiber blended with charged surface modifying macromolecules (cSMM) was successfully fabricated. This study concludes that PEGHBS and DEGHBS were successfully synthesized and having an end-capped that able to induce charge at membrane surface. From the modeling sodium chloride rejection mechanism, it was revealed that the sulfonate properties were found slightly higher charge than the carboxylate group as studied in this work.

ACKNOWLEDGEMENT

The authors would like to acknowledge the financial support from Ministry of Higher Education (MOHE) Malaysia. The author is thankful to both Universiti Teknologi Malaysia and Universiti Malaysia Sabah for providing excellent research environment in which to complete this work.

REFERENCES

[1]. Rana, D., T. Matsuura, R.M. Narbaitz, C. Feng, Development and characterization of novel hydrophilic surface modifying macromolecule for polymeric membranes, *J. Membr. Sci.* 249:103. (2005)

[2]. Rachwal A.J, Khow J, Colbourne J, S.O'Donnel, Water Treatment For Public Supply In The 1990's – A Role For Membrane Technology? *Desalination*, 427-436. (1994)

[3]. Cartwright P.S, Industrial Wastewater Treatment with Membranes – US Perspective. Water Sci. Technology 25 (10), 373-390. (1992)

[4]. Choi, J.H, Fukushi K., Yamamoto K, (2006). A Submerged Nanofiltration Membrane Bioreactor for Domestic Wastewater Treatment: the Performance of CA Nanofiltration Membranes for Long-Term Operation, separation and purification technology, 52-3:470.

[5]. Choi, J.H, Dockko S, Fukushi K, Yamamoto K, (2002). A Novel Application of a Submerged Nanofiltration Membrane Bioreactor (NF MBR) for Wastewater Treatment, Desalination 146, 413.

[6]. Clara, M., B. Strenn, O. Gans, E. Martinez, N. Kreuzinger, H. Kroiss, (2005). Removal of Selected Pharmaceuticals, Fragrances and Endocrine Disrupting Compounds in a Membrane Bioreactor and Conventional Wastewater Treatment Plants, *Water Research* 39:4797.

[7]. Schafer, A.I, (2001). *Natural Organics Removal Using Membranes: Principles, Performance and Cost*, Technomic Pulishing co. Inc, USA.

[8]. Bellona C., J.E Drewes, P. Xu, G. Amy, Factors affecting the rejection of organic solutes during NF/RO treatment—a literature review, *Water Research* 38: 2795–2809. (2004)

[9]. Braeken L., B. Bettens, K. Boussu, P. Van De Meeren, J. Cocquyt, J. Vermant, B. Van der Bruggen, Transport Mechanism Of Dissolved Organic Compound In Aqueous Solution During Nanofiltration, *J. Membr. Sci.*, 279 (1-2):311-319. (2006).

[10]. Bruggen B V., J. Schaep, D. Williams, C. Vandecasteele, Influence of molecular size, polarity and charge on the retention of organic molecules by nanofiltration, *J. Membr. Sci.* 156,29-41. (1999).

[11]. Verliefde A.R.D., S.G.J. Heijman, E.R. Cornelissen, G. Amy, B. Van der Bruggen, J.C. van Dijk, Influence of electrostatic interactions on the rejection with NF and assessment of the removal efficiency during NF/GAC treatment of pharmaceutically active compounds in surface water, *Water Research* 41, 3227-3240. (2007)

[12]. Bellona C. & J.E Drewes, The role of membrane surface charge and solute physico-chemical properties in the rejection of organic acids by NF membranes, *J. Membr. Sci*. 249, 227-234. (2005).

[13]. Tanninen J., M. Nystrom, separation of ions in acidic conditions using NF, *Desalination*, 147:295-299. (2002).

[14]. Childress A.E, M. Elimelechrelating nanofiltration membrane performance to membrane charge (electrokinetic) characteristics, *Environ. Sci. Technol*; 34: 3710-3716. (2000).

[15]. Yoon J., G. Amy, J. Cho, N. Her, J. Pellergrino, Transport of perchlorate (ClO^-) through NF and UF membranes, *Desalination* 147: 11-17. (2002).

[16]. Shim Y., H-G Lee, S. Lee, S-H Moon, J. Cho, Effects of NOM and ionic species on membrane surface charge, Environ. Sci. Technol. 36: 3864-3871. (2002).

[17]. Verliefde A., E. Cornelissen, G. Amy, B. Van der Bruggen, Hans van Dijk, Priority Organic Micropollutants In Water Sources In Flanders And The Netherlands And Assessment Of Removal Possibilities With Nanofiltration, *Environmental Pollution* 146: 281-289. (2007).

[18]. Bolong, N., Ismail A.F., Salim M.R., D. Rana and T. Matsuura. Development and Characterization of Novel Charged Surface Modification Macromolecule to Polyethersulfone Hollow Fiber Membrane with Polyvinylpyrrolidone and Water, *J. Membrane Science*, 331: 40. (2009).

[19]. Spiegler, K.S. and Kedem O., Thermodynamics of Hyperfiltration (Reverse Osmosis): Criteria for Efficient Membranes, *Desalination.* 1:311-326. (1966).

[20]. McKetta, J. J., G. E. Wiesmantel, *Encyclopedia of Chemical Processing and Design*, CRC Press, Florida. (1997)

[21]. Bowen WR, Mohammad A.W., Characterization and prediction of nanofiltration membrane performance – a general assessment. Trans Inst Chem Eng 76A:885–893. (1998).

[22]. Hassan, A. R., Ali N., Abdullah N., Ismail A.F., A theoretical approach on membrane characterization: the deduction of fine structural details of asymmetric nanofiltration membranes, *Desalination*. 206: 107. (2007).

[23]. Kedem, O., A. Katchalsky, Permeability of composite membranes, *Trans. Faraday Soc.* 59:1918. (1963).

[24]. Spiegler, K.S. and Kedem O., Thermodynamics of hyperfiltration (reverse osmosis): criteria for efficient membranes, *Desalination.* 1:311-326. (1966).

[25]. Ismail, A. F., Hassan A. R., Effect of additive contents on the performances and structural properties of asymmetric polyethersulfone (PES) nanofiltration membranes, *Separation and. Purification Tech*. 55:98. (2007).

Design of an Adaptive Neural Predictive Nonlinear Controller for Nonholonomic Mobile Robot System Based on Posture Identifier in the Presence of Disturbance

Ahmed S. Al-Araji, Maysam F. Abbod and Hamed S. Al-Raweshidy
Wireless Networks and Communication Centre, School of Engineering and Design
Brunel University
London - UK
ahmedsas2040@yahoo.com

***Abstract* —This paper proposes an adaptive neural predictive nonlinear controller to guide a nonholonomic wheeled mobile robot during continuous and non-continuous gradients trajectory tracking. The structure of the controller consists of two models that describe the kinematics and dynamics of the mobile robot system and a feedforward neural controller. The models are modified Elman neural network and feedforward multi-layer perceptron respectively. The modified Elman neural network model is trained off-line and on-line stages to guarantee the outputs of the model accurately represent the actual outputs of the mobile robot system. The trained neural model acts as the position and orientation identifier. The feedforward neural controller is trained off-line and adaptive weights are adapted on-line to find the reference torques, which controls the steady-state outputs of the mobile robot system. The feedback neural controller is based on the posture neural identifier and quadratic performance index optimization algorithm to find the optimal torque action in the transient state for N-step-ahead prediction. General back propagation algorithm is used to learn the feedforward neural controller and the posture neural identifier. Simulation results show the effectiveness of the proposed adaptive neural predictive control algorithm; this is demonstrated by the minimised tracking error and the smoothness of the torque control signal obtained with bounded external disturbances.**

Keywords - Nonholonomic Mobile Robots, Adaptive Predictive Nonlinear Controller, Neural Networks, Trajectory Tracking.

I. INTRODUCTION

In recent years, wheel-based mobile robots have attracted considerable attention in various industrial and service applications. For example, room cleaning, factory automation, transportation, etc. These applications require mobile robots to have the ability to track specified path stably [1]. In general, nonholonomic behaviour in robotic systems is particularly interesting because this mechanism can completely be controlled with reduced number of actuators. Several controllers were proposed for trajectory tracking of mobile robots with nonholonomic constraints. The traditional control methods for mobile robot path tracking have used linear or non-linear feedback control while artificial intelligent controllers were carried out using neural networks or fuzzy inference [2].

There are other techniques for trajectory tracking controllers such as predictive control technique. Predictive approaches to path tracking seem to be very promising because the reference trajectory is known beforehand. Model predictive trajectory tracking control was applied to a mobile robot where linearised tracking error dynamics was used to predict future system behaviour and a control law was derived from a quadratic cost function penalizing the system trucking error and the control effort [3].

In addition, an adaptive trajectory-tracking controller based on the robot dynamics was proposed in [4 and 5] and its stability property was proved using the Lyapunov theory. An adaptive controller of nonlinear PID-based neural networks was developed for the velocity and orientation tracking control of a nonholonomic mobile robot [6].

A trajectory tracking control for a nonholonomic mobile robot by the integration of a kinematics controller and neural dynamic controller based on the sliding mode theory was presented in [7]. The adaptive feedforward and feedback neural controllers with predictive optimization algorithm have minimised the tracking error of the nonholonomic wheeled mobile robot as presented in [8].

Two novel dual adaptive neural control schemes were proposed for dynamic control of nonholonomic mobile robots [9]. The first scheme was based on Gaussian radial basis function ANNs and the second on sigmoidal multilayer perceptron (MLP) ANNs. ANNs were employed for real-time approximation of the robot's nonlinear dynamic functions which were assumed to be unknown. Integrating the neural networks into back-stepping technique has improved learning algorithm of analogue compound orthogonal networks and novel tracking control approach for nonholonomic mobile robots [10]. A variable structure control algorithm was proposed to study the trajectory tracking control based on the kinematics model of a 2-wheel differentially driven mobile robot by using of the back stepping method and virtual feedback parameter with the sigmoid function [11]. The trajectory-tracking controllers designed by pole-assignment approach for mobile robot model were presented in [12].

The contribution of the presented approach is the analytically derived control law which has significantly high

computational accuracy with predictive optimization technique to obtain the optimal torques control action and lead to minimum tracking error of the mobile robot for different types of trajectories with continuous gradients such as (lemniscates) or non-continuous gradients (square) with bounded external disturbances.

The predictive optimization algorithm for N step ahead can generate excellent feedback control action in order to reduce the effect of external disturbances.

Simulation results show that the proposed controller is robust and effective in terms of fast response and minimum tracking error and in generating an optimal torque control action despite of the presence of bounded external disturbances.

The remainder of the paper is organized as follows. Section two is a description of the kinematics and dynamics model of the nonholonomic wheeled mobile robot. In section three, the proposed adaptive neural predictive controller is derived. The simulation results of the proposed controller are presented in section four and the conclusions are drawn in section five.

II. The Kinematics and Dynamics Model of Nonholonomic Wheeled Mobile Robot

The schematic of the nonholonomic mobile robot, shown in figure 1, consists of a cart with two driving wheels mounted on the same axis and an omni-directional castor in the front of cart. The castor carries the mechanical structure and keeps the platform more stable [6 and 8]. Two independent analogous DC motors are the actuators of left and right wheels for motion and orientation. The two wheels have the same radius denoted by r, and L is the distance between the two wheels. The centre of mass of the mobile robot is located at point c, centre of axis of wheels.

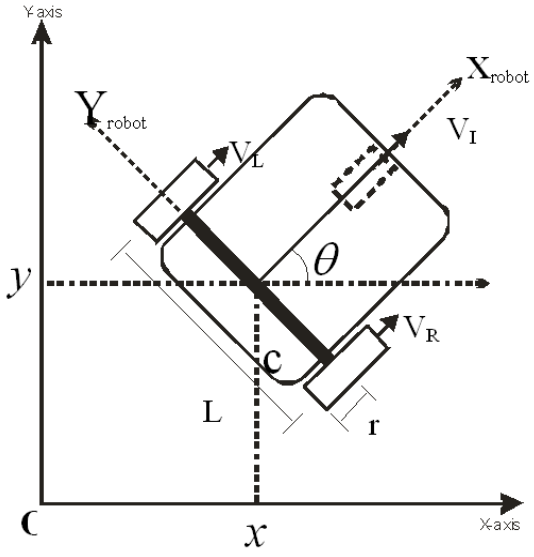


Figure 1. Schematic of the nonholonomic mobile robot.

The pose of mobile robot in the global coordinate frame $[O, X, Y]$ and the pose vector in the surface is defined as:

$$q = (x, y, \theta)^T \quad (1)$$

where $q(t) \in \Re^{3\times1}$,

x and y are coordinates of point c and θ is the robotic orientation angle measured with respect to the X-axis. These three generalized coordinates can describe the configuration of the mobile robot. The mobile robot is subjected to an independent velocity constraint that can be expressed in matrix form [13]:

$$A^T(q)\dot{q} = 0 \quad (2)$$

where

$$A^T(q) = [-\sin\theta(t) \quad \cos\theta(t) \quad 0] \quad (3)$$

$A(q) \in \Re^{3\times1}$

It is assumed that the mobile robot wheels are ideally installed in such a way that they have ideal rolling without skidding [14].

Therefore, the kinematics of the robot can be described as

$$\dot{q} = \begin{bmatrix} \dot{x}(t) \\ \dot{y}(t) \\ \dot{\theta}(t) \end{bmatrix} = \begin{bmatrix} \cos\theta(t) & 0 \\ \sin\theta(t) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} V_I(t) \\ V_w(t) \end{bmatrix} \quad (4)$$

where $S(q)$ is defining a full rank matrix as

$$S(q) = \begin{bmatrix} \cos\theta(t) & 0 \\ \sin\theta(t) & 0 \\ 0 & 1 \end{bmatrix} \quad (5)$$

where V_I and V_w, the linear and angular velocities.

Forces must be applied to the mobile robot to produce motion. These forces are modeled by studying the motion of the dynamic model of the differential wheeled mobile robot shown in figure 1. Mass, forces and speed are associated with this motion. The dynamic model can be described by the following form of dynamic equations based on Euler Lagrange formulation [5, 6, 8 and 9].

$$M(q)\ddot{q} + C(q,\dot{q})\dot{q} + G(q) + \tau d = B(q)\tau - A^T(q)\lambda \quad (6)$$

$M(q) \in \Re^{3\times3}$ is a symmetric positive definite inertia matrix, $C(q,\dot{q}) \in \Re^{3\times3}$ is the centripetal and carioles matrix, $G(q) \in \Re^{3\times1}$ is the gravitational torques vector, $\tau d \in \Re^{3\times1}$ denotes bounded unknown disturbances including unstructured and unmodeled dynamics, $B(q) \in \Re^{3\times2}$ is the input transformation matrix, $\tau \in \Re^{2\times1}$ is input torque vector, and $\lambda \in \Re^{1\times1}$ is the vector of constraint forces.

Remark 1: The plane of each wheel is perpendicular to the ground and the contact between the wheels and the ground is pure rolling and non-slipping, and hence the velocity of the centre of the mass of the mobile robot is orthogonal to the rear wheels' axis.

Remark 2: The trajectory of mobile robot base is constrained to the horizontal plane, therefore, $G(q)$ is equal to zero.

Remark 3: In this dynamic model, the passive self-adjusted supporting wheel influence is not taken into

consideration as it is a free wheel. This significantly reduces the complexity of the model for the feedback controller design. However, the free wheel may be a source of substantial distortion, particularly in the case of changing its movement direction. This effect is reduced if the small velocity of the robot is considered [5 and 6]. Remark 4: The centre of mass for mobile robot is located in the middle of axis connecting the rear wheels in C point as shown in figure 1, therefore, $C(q,\dot{q})$ is equal to zero.

The dynamical equation of the differential wheeled mobile robot can be expressed as

$$\begin{bmatrix} M & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} \ddot{x} \\ \ddot{y} \\ \ddot{\theta} \end{bmatrix} + \tau d = \frac{1}{r} \begin{bmatrix} \cos\theta & \cos\theta \\ \sin\theta & \sin\theta \\ \frac{L}{2} & \frac{-L}{2} \end{bmatrix} \begin{bmatrix} \tau_L \\ \tau_R \end{bmatrix} + \begin{bmatrix} -\sin\theta \\ \cos\theta \\ 0 \end{bmatrix} \lambda \tag{7}$$

where τ_L and τ_R are the torques of left and right motors respectively. M and I present the mass and inertia of the mobile robot respectively.

By solving equation (4 and 7) then we can reach the normal form,

$$\dot{V}_I = \frac{\tau_L + \tau_R}{Mr} + \tau d \tag{8}$$

$$\dot{V}_W = \frac{L(\tau_L - \tau_R)}{2rI} + \tau d \tag{9}$$

where $\dot{V}_I$ and $\dot{V}_W$ are the linear and angular acceleration of the differential wheeled mobile robot.

The dynamics and the kinematics model structure of the differential wheeled mobile robot can be shown in figure 2.

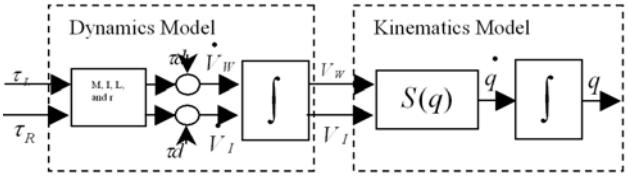


Figure 2. Dynamics and kinematics model structure of the mobile robot.

III. Adaptive Neural Predictive Control Methodology

The control of nonlinear MIMO mobile robot system is considered in this section. The approach to control the mobile robot depends on the available information of the unknown nonlinear system can be known by the input-output data only and the control objectives. The first step in the procedure of the control structure is the identification of the kinematics and dynamics mobile robot from the input-output data. Then an adaptive feedforward neural controller is designed to find reference torques that control the steady-state outputs of the mobile robot trajectory.

The feedback neural controller is based on the minimisation of a quadratic performance index function of the error between the desired trajectory input and the posture neural identifier output, i.e. position and orientation of mobile robot trajectory, and the feedback neural controller itself. The predictive optimization algorithm is used to determine the torque control signal for N-steps-ahead and to use minimum torque effort. The torque control signal will minimise the cost function in order to minimise the tracking error as well as reduce the torque control effort in the presence of external disturbance. The integrated adaptive control structure, which consists of an adaptive feedforward neural controller and feedback neural controller with an optimization algorithm, brings together the advantages of the adaptive neural method with the robustness of feedback for N-step-ahead prediction.

The proposed structure of the adaptive neural predictive controller can be given in the form of block diagram as shown in figure 3. It consists of:

a) Position and Orientation Neural Networks Identifier.
b) Feedforward Neural Controller.
c) Feedback Neural Controller.

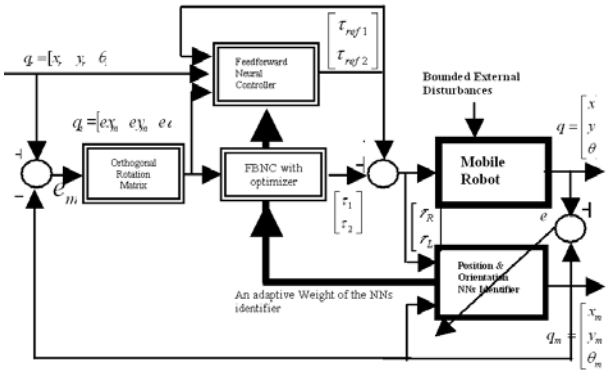


Figure. 3. The proposed structure of the adaptive neural predictive controller for the nonholonomic wheeled mobile robot.

A. Position and Orientation Neural Networks Identifier

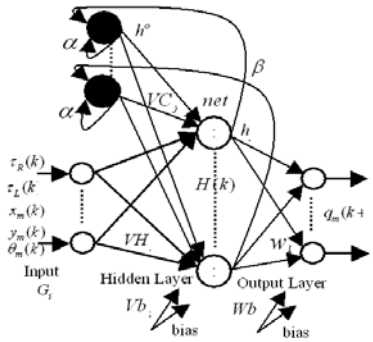


Figure 4 . Elman neural networks acts as the posture identifier.

Nonlinear MIMO system identification of kinematics and dynamics mobile robot, position and orientation, will be introduced in this section. The modified Elman recurrent neural network model is applied to construct the position and orientation neural network identifier as shown in figure

4. The nodes of input, context, hidden and output layers are highlighted. The network uses two configuration models, series-parallel and parallel identification structures, which are trained using dynamic back-propagation algorithm.

The structure shown in figure 4 is based on the following equations [15]:

$$h(k) = F\{VH\,\overline{G}(k), VC\,\overline{h^o}(k), bias\overline{Vb}\} \quad (10)$$

$$O(k) = (Wh(k), bias\overline{Wb}) \quad (11)$$

where *VH*,*VC* and *W* are weight matrices, $\overline{Vb}$ and $\overline{Wb}$ are weight vectors and *F* is a non-linear vector function. The multi-layered modified Elman neural network, shown in figure 4, is composed of many interconnected processing units called neurons or nodes.

The output of the context unit in the modified Elman network is given by [15]:

$$h_c^o(k) = \alpha h_c^o(k-1) + \beta h_c(k-1) \quad (12)$$

where $h_c^o(k)$ and $h_c(k)$ are the outputs of the context and hidden units respectively. α is the feedback gain of the self-connections and β is the connection weight from the hidden units (j^{th}) to the context units (c^{th}) at the context layer. The value of α and β are selected randomly between (0 and 1) [15].

The outputs of the identifier are the modelling pose vector in the surface and are defined as:

$q_m = (x_m, y_m, \theta_m)^T$, where x_m and y_m are the modelling coordinates and θ_m is the modelling orientation angle.

The learning algorithm will be used to adjust the weights of dynamical recurrent neural network. Dynamic back propagation algorithm is used to train the Elman network. The sum of the square of the differences between the desired outputs $q = (x, y, \theta)^T$ and neural network identifier outputs $q_m = (x_m, y_m, \theta_m)^T$ is given by equation (13).

$$E = \frac{1}{2}\sum_{i=1}^{np}((x-x_m)^2 + (y-y_m)^2 + (\theta-\theta_m)^2) \quad (13)$$

where *np* is the number of patterns.

The connection matrix between hidden layer and output layer is W_{kj}

$$\Delta W_{kj}(k+1) = -\eta\frac{\partial E}{\partial W_{kj}} \quad (14)$$

where η is learning rate.

$$\frac{\partial E}{\partial W_{kj}} = \frac{\partial E}{\partial q_m(k+1)}\frac{\partial q_m(k+1)}{\partial o_k}\frac{\partial o_k}{\partial net_k}\frac{\partial net_k}{\partial W_{kj}} \quad (15)$$

$$W_{kj}(k+1) = W_{kj}(k) + \Delta W_{kj}(k+1) \quad (16)$$

The connection matrix between input layer and hidden layer is VH_{ji}

$$\Delta VH_{ji}(k+1) = -\eta\frac{\partial E}{\partial VH_{ji}} \quad (17)$$

$$\frac{\partial E}{\partial VH_{ji}} = \frac{\partial E}{\partial q_m(k+1)}\frac{\partial q_m(k+1)}{\partial o_k}\frac{\partial o_k}{\partial net_k}\frac{\partial net_k}{\partial h_j}\frac{\partial h_j}{\partial net_j}\frac{\partial net_j}{\partial VH_{ji}} \quad (18)$$

$$VH_{ji}(k+1) = VH_{ji}(k) + \Delta VH_{ji}(k+1) \quad (19)$$

The connection matrix between context layer and hidden layer is VC_{ji}

$$\Delta VC_{jc}(k+1) = -\eta\frac{\partial E}{\partial VC_{jc}} \quad (20)$$

$$\frac{\partial E}{\partial VC_{jc}} = \frac{\partial E}{\partial q_m(k+1)}\frac{\partial q_m(k+1)}{\partial o_k}\frac{\partial o_k}{\partial net_k}\frac{\partial net_k}{\partial h_j}\frac{\partial h_j}{\partial net_c}\frac{\partial net_c}{\partial VC_{jc}} \quad (21)$$

$$VC_{jc}(k+1) = VC_{jc}(k) + \Delta VC_{jc}(k+1) \quad (22)$$

B. Feedforward Neural Controller

The feedforward neural controller (FFNC) is of prime importance in the structure of the controller due to its necessity in keeping the steady-state tracking error at zero. This means that the actions of the FFNC, $\tau_{ref1}(k)$ and $\tau_{ref2}(k)$ are used as the reference torques of the steady state outputs of the mobile robot. Hence, the FFNC is supposed to learn the adaptive inverse model of the mobile robot system with off-line and on-line stages to calculate mobile robot's reference input torques drive. Reference input torques will keep the robot on a desired trajectory in the presence of any disturbances or initial state errors. To achieve FFNC, a multi-layer perceptron model is used as shown in figure 5 [16].

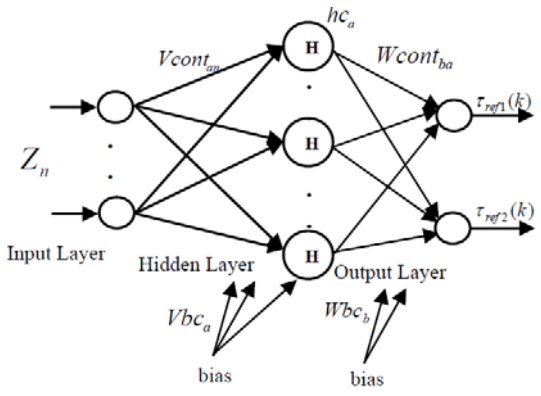


Figure 5. MLP Neural network acts as the feed forward neural controller

The training of the feedforward neural controller is performed off-line as shown in figure 6, which the weights adapted on-line. It depends on the posture neural network identifier to find the mobile robot Jacobian through the neural identifier model.

This approach is currently considered as one of the better approaches that can be followed to overcome the lack of initial knowledge. The dynamic back propagation algorithm is employed to realize the training the weights of the feedforward neural controller.

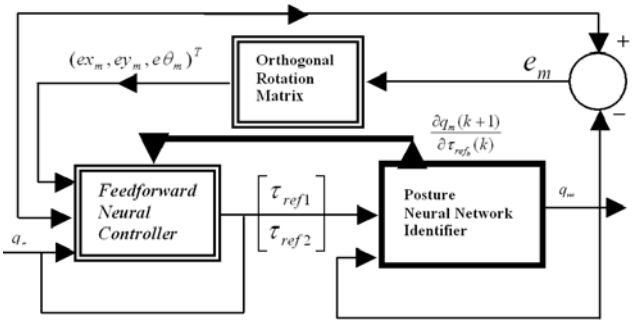


Figure 6. The feedforward neural controller structure for mobile robot model.

The sum of the square of the differences between the desired posture $q_r = (x_r, y_r, \theta_r)^T$ and neural network posture $q_m = (x_m, y_m, \theta_m)^T$ is:

$$Ec = \frac{1}{2}\sum_{i=1}^{npc}((x_r - x_m)^2 + (y_r - y_m)^2 + (\theta_r - \theta_m)^2) \quad (23)$$

where *npc* is number of patterns.

The connection matrix between hidden layer and output layer is $Wcont_{ba}$

$$\Delta Wcont_{ba}(k+1) = -\eta \frac{\partial Ec}{\partial Wcont_{ba}} \quad (24)$$

$$\frac{\partial Ec}{\partial Wcont_{ba}} = \frac{\partial Ec}{\partial q_m(k+1)} \frac{\partial q_m(k+1)}{\partial \tau_{ref_b}(k)} \frac{\partial \tau_{ref_b}(k)}{\partial oc_b} \frac{\partial oc_b}{\partial netc_b} \frac{\partial netc_b}{\partial Wcont_{ba}} \quad (25)$$

$$\frac{\partial Ec}{\partial q_m(k+1)} = \frac{\partial \frac{1}{2}\sum((x_r - x_m)^2 + (y_r - y_m)^2 + (\theta_r - \theta_m)^2)}{\partial q_m(k+1)} \quad (26)$$

This is achieved in the local coordinates with respect to the body of the mobile robot that is the same outputs of the position and orientation neural networks identifier. The configuration error can be represented by using a transformation matrix as:

$$\begin{bmatrix} ex_m \\ ey_m \\ e\theta_m \end{bmatrix} = \begin{bmatrix} \cos\theta_m & \sin\theta_m & 0 \\ -\sin\theta_m & \cos\theta_m & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_r - x_m \\ y_r - y_m \\ \theta_r - \theta_m \end{bmatrix} \quad (27)$$

where x_r, y_r and θ_r are the reference posture of the mobile robot.

$$Jacobian = \frac{\partial q_m(k+1)}{\partial \tau_{ref_b}(k)} \quad (28)$$

where the outputs of the identifier are $q_m = (x_m, y_m, \theta_m)^T$.

$$\frac{\partial q_m(k+1)}{\partial \tau_{ref_b}(k)} = \frac{\partial q_m(k+1)}{\partial o_k(k)} \frac{\partial o_k(k)}{\partial net_k} \frac{\partial net_k}{\partial h_j} \frac{\partial h_j}{\partial net_j} \frac{\partial net_j}{\partial \tau_{ref_b}(k)} \quad (29)$$

Substituting equations (26 and 29) into equation (25), to find $\Delta Wcont_{ba}(k+1)$, then

$$Wcont_{ba}(k+1) = Wcont_{ba}(k) + \Delta Wcont_{ba}(k+1) \quad (30)$$

The connection matrix between input layer and hidden layer is $Vcont_{an}$

$$\Delta Vcont_{an}(k+1) = -\eta \frac{\partial Ec}{\partial Vcont_{an}} \quad (31)$$

$$\frac{\partial Ec}{\partial Vcont_{an}} = \frac{\partial Ec}{\partial q_m(k+1)} \frac{\partial q_m(k+1)}{\partial \tau_{ref_b}(k)} \frac{\partial \tau_{ref_b}(k)}{\partial oc_b} \times$$

$$\frac{\partial oc_b}{\partial netc_b} \times \frac{\partial netc_b}{\partial hc_a} \times \frac{\partial hc_a}{\partial netc_a} \times \frac{\partial netc_a}{\partial Vcont_{an}} \quad (32)$$

Substituting equations (26 and 29) into equation (32), to find $\Delta Vcont_{an}(k+1)$, then

$$Vcont_{an}(k+1) = Vcont_{an}(k) + \Delta Vcont_{an}(k+1) \quad (33)$$

Once the feedforward neural controller has learned, it generates the torque control action to keep the output of the mobile robot at the steady state reference value and to overcome any external disturbances during trajectory.

The torques will be known equivalently as τ_{ref1} and τ_{ref2}, the reference torques of the right and left wheels respectively.

C. *Feedback Neural Controller*

The feedback neural controller is essential to stabilise the tracking error of the mobile robot system when the trajectory of the robot is drifted from the reference trajectory during transient state. The feedback neural controller generates an optimal torque control action that minimises the cumulative error between the reference input trajectory and the output trajectory of the mobile robot. The weighted sum of the torque control signal can be obtained by minimising a quadratic performance index. The feedback neural controller consists of the adaptive weights of the position and orientation neural networks identifier and an optimization algorithm. The quadratic performance index for multi input /multi output system can be expressed as:

$$J = \frac{1}{2}\sum_{k=1}^{N} Q(q_r(k+1) - q(k+1))^2 + R((\tau_{ref1}(k) - \tau_R(k))^2 + (\tau_{ref2}(k) - \tau_L(k))^2) \quad (34)$$

Hence

$$q_r(k+1) = [x_r(k+1), y_r(k+1), \theta_r(k+1)]^T \quad (35)$$

$$q(k+1) = [x(k+1), y(k+1), \theta(k+1)]^T \quad (36)$$

$$\tau_R(k) = \tau_{ref1}(k) + \tau_1(k) \quad (37)$$

$$\tau_L(k) = \tau_{ref2}(k) + \tau_2(k) \quad (38)$$

(Q, R) are positive weighting factors.
N is the number of steps ahead.

The quadratic cost function will not only force the mobile robot output to follow the reference trajectory by minimising the cumulative error for N steps ahead but also forces the torque control actions ($\tau_1(k)$ and $\tau_2(k)$) in the transient period to be as close as possible to the reference torque control signals ($\tau_{ref1}(k)$ and $\tau_{ref2}(k)$). In addition, J depends on Q & R factors and chooses a set of values of the weighting factors Q and R to determine the optimal control action by observing the system behavior [17]. The on-line position and orientation neural networks identifier is to be used to obtain the predicted values of the outputs of the mobile robot system $q_m(k+1)$ for N steps ahead instead of running the mobile robot system itself $q(k+1)$ for N steps. This is performed to find the optimal torque control actions by using the posture identifier weights and optimization algorithm depending on the quadratic cost function. Therefore, it can be said that:

$$q_m(k+1) \approx q(k+1) \tag{39}$$

The performance index of equation (34) can be put as:

$$J = \frac{1}{2}\sum_{k=1}^{N} Q((x_r(k+1)-x_m(k+1))^2 + (y_r(k+1)-y_m(k+1))^2 + (\theta_r(k+1)-\theta_m(k+1))^2) + R((\tau_1(k))^2 + (\tau_2(k))^2) \tag{40}$$

To achieve equation (39 and 40), the modified Elman neural network will be used as posture identifier. This task is carried out using an identification technique based on series-parallel and parallel configuration with two stages to learn the posture identifier. The first stage is an off-line identification, while the second stage is an on-line modification of the weights of the obtained position and orientation neural identifier. The on-line modifications are necessary to keep tracking any possible variation in the kinematics and dynamics parameters of the mobile robot system.

Back propagation algorithm (BPA) is used to adjust the weights of the posture neural identifier to learn the kinematics and dynamics of the mobile robot system, by applying a simple gradient decent rule.

For N steps estimation of the two feedback neural controller actions $\tau_1(k)\,\&\,\tau_2(k)$ the techniques of generalized predictive control theory will be used. The N steps estimation of $\tau_1(k)\,\&\,\tau_2(k)$ will be calculated for each sample. The position and orientation in the identifier model, shown in figure 4, represent the kinematics and dynamics model of the mobile robot system and will be controlled asymptotically. Therefore, they can be used to predict future values of the model outputs for the next N steps and can be used to find the optimal value of $\tau_1(k)\,\&\,\tau_2(k)$ using an optimization algorithm.

For this purpose, let N be a pre-specified positive integer that is denoted such that the future values of the set point are:

$$X_{rt,N} = [x_r(t+1), x_r(t+2), x_r(t+3), ..., x_r(t+N)] \tag{41}$$

$$Y_{rt,N} = [y_r(t+1), y_r(t+2), y_r(t+3), ..., y_r(t+N)] \tag{42}$$

$$\theta_{rt,N} = [\theta_r(t+1), \theta_r(t+2), \theta_r(t+3), ..., \theta_r(t+N)] \tag{43}$$

As the future values of set point and (t) represents the time instant, and the predicted outputs of the robot model used the neural identifier, shown in figure 4, are:

$$X_{mt,N} = [x_m(t+1), x_m(t+2), x_m(t+3), ..., x_m(t+N)] \tag{44}$$

$$Y_{mt,N} = [y_m(t+1), y_m(t+2), y_m(t+3), ..., y_m(t+N)] \tag{45}$$

$$\theta_{mt,N} = [\theta_m(t+1), \theta_m(t+2), \theta_m(t+3), ..., \theta_m(t+N)] \tag{46}$$

The error vector of position and orientation as equations (47, 48, and 49) can be calculated by using equation (27).

$$EX_{m,t,N} = [ex_m(t+1), ex_m(t+2), ex_m(t+3)..., ex_m(t+N] \tag{47}$$

$$EY_{m,t,N} = [ey_m(t+1), ey_m(t+2), ey\ (t+3)..., ey_m(t+N] \tag{48}$$

$$E\theta_{m,t,N} = [e\theta_m(t+1), e\theta_m(t+2), e\theta_m(t+3)..., e\theta_m(t+N] \tag{49}$$

Two-feedback control signals can be determined by:

$$\tau'_{1t,N} = [\tau'_1(t), \tau'_1(t+1), \tau'_1(t+2), ..., \tau'_1(t+N-1)] \tag{50}$$

$$\tau'_{2t,N} = [\tau'_2(t), \tau'_2(t+1), \tau'_2(t+2), ..., \tau'_2(t+N-1)] \tag{51}$$

Assuming the following objective function:

$$J1 = \frac{1}{2}Q[(EX_{m,t,N}EX_{m,t,N}{}^T) + (EY_{m,t,N}EY_{m,t,N}{}^T) + (E\theta_{m,t,N}E\theta_{m,t,N}{}^T)] + \frac{1}{2}R[(\tau'_{1t,N}\tau'_{1t,N}{}^T) + (\tau'_{2t,N}\tau'_{2t,N}{}^T)] \tag{52}$$

then it is aimed to find τ'_1 and τ'_2 such that J1 is minimised using the gradient descent rule. The new control actions will be given by:

$$\tau'_{1t,N}{}^{K+1} = \tau'_{1t,N}{}^{K} + \Delta\tau'_{1t,N}{}^{K} \tag{53}$$

$$\tau'_{2t,N}{}^{K+1} = \tau'_{2t,N}{}^{K} + \Delta\tau'_{2t,N}{}^{K} \tag{54}$$

where *k* here indicates that calculations are performed at the k^{th} sample; and

$$\Delta\tau'_{1t,N}{}^{K} = -\eta\frac{\partial J1}{\partial\tau'_{1t,N}{}^{K}} = [\Delta\tau'_1(t), \Delta\tau'_1(t+1), \Delta\tau'_1(t+2), ...\Delta\tau'_1(t+N-1)] \tag{55}$$

$$\Delta\tau'_{2t,N}{}^{K} = -\eta\frac{\partial J1}{\partial\tau'_{2t,N}{}^{K}} = [\Delta\tau'_2(t), \Delta\tau'_2(t+1), \Delta\tau'_2(t+2), ...\Delta\tau'_2(t+N-1)] \tag{56}$$

$$-\eta\frac{\partial J1}{\partial\tau'_{1t,N}{}^{K}} = \eta QEX_{m,t,N}\frac{\partial X_{mt,N}}{\partial\tau'_{1t,N}{}^{K}} + \eta QEY_{m,t,N}\frac{\partial Y_{mt,N}}{\partial\tau'_{1t,N}{}^{K}} + \eta QE\theta_{m,t,N}\frac{\partial\theta_{mt,N}}{\partial\tau'_{1t,N}{}^{K}} - \eta R\partial\tau'_{1t,N}{}^{K} \tag{57}$$

$$-\eta\frac{\partial J1}{\partial\tau'_{2t,N}{}^{K}} = \eta QEX_{m,t,N}\frac{\partial X_{mt,N}}{\partial\tau'_{2t,N}{}^{K}} + \eta QEY_{m,t,N}\frac{\partial Y_{mt,N}}{\partial\tau'_{2t,N}{}^{K}} + \eta QE\theta_{m,t,N}\frac{\partial\theta_{mt,N}}{\partial\tau'_{2t,N}{}^{K}} - \eta R\partial\tau'_{2t,N}{}^{K} \tag{58}$$

Equations (59 to 64) are the well-known Jacobian vectors.

$$\frac{\partial X_{mt,N}}{\partial\tau'_{1t,N}{}^{K}} = \left[\frac{\partial x_m(t+1)}{\partial\tau'_1(t)}\ \ \frac{\partial x_m(t+2)}{\partial\tau'_1(t+1)}\ \ ...\ \ \frac{\partial x_m(t+N)}{\partial\tau'_1(t+N-1)}\right] \tag{59}$$

$$\frac{\partial X_{mt,N}}{\partial\tau'_{2t,N}{}^{K}} = \left[\frac{\partial x_m(t+1)}{\partial\tau'_2(t)}\ \ \frac{\partial x_m(t+2)}{\partial\tau'_2(t+1)}\ \ ...\ \ \frac{\partial x_m(t+N)}{\partial\tau'_2(t+N-1)}\right] \tag{60}$$

$$\frac{\partial Y_{mt,N}}{\partial\tau'_{1t,N}{}^{K}} = \left[\frac{\partial y_m(t+1)}{\partial\tau'_1(t)}\ \ \frac{\partial y_m(t+2)}{\partial\tau'_1(t+1)}\ \ ...\ \ \frac{\partial y_m(t+N)}{\partial\tau'_1(t+N-1)}\right] \tag{61}$$

$$\frac{\partial Y_{mt,N}}{\partial\tau'_{2t,N}{}^{K}} = \left[\frac{\partial y_m(t+1)}{\partial\tau'_2(t)}\ \ \frac{\partial y_m(t+2)}{\partial\tau'_2(t+1)}\ \ ...\ \ \frac{\partial y_m(t+N)}{\partial\tau'_2(t+N-1)}\right] \tag{62}$$

$$\frac{\partial\theta_{mt,N}}{\partial\tau'_{1t,N}{}^{K}} = \left[\frac{\partial\theta_m(t+1)}{\partial\tau'_1(t)}\ \ \frac{\partial\theta_m(t+2)}{\partial\tau'_1(t+1)}\ \ ...\ \ \frac{\partial\theta_m(t+N)}{\partial\tau'_1(t+N-1)}\right] \tag{63}$$

$$\frac{\partial\theta_{mt,N}}{\partial\tau'_{2t,N}{}^{K}} = \left[\frac{\partial\theta_m(t+1)}{\partial\tau'_2(t)}\ \ \frac{\partial\theta_m(t+2)}{\partial\tau'_2(t+1)}\ \ ...\ \ \frac{\partial\theta_m(t+N)}{\partial\tau'_2(t+N-1)}\right] \tag{64}$$

It can be seen that each element in the above vectors can be calculated from equation (65 to 74) such that:

$$net_j = \sum_{i=1}^{nh} VH_{ji} \times G_i + \sum_{c=1}^{C} VC_{jc} \times h_c^o + bias \times Vb_j \tag{65}$$

where $j=c$ and $nh=C$ are the number of the hidden and context nodes respectively and $\overline{G}$ is the input vector such as

$$G = [\tau_R(t), \tau_L(t), x_m(t), y_m(t), \theta_m(t)]^T \tag{66}$$

$$h_j = \frac{2}{1+e^{-net_j}} - 1 \tag{67}$$

$$f(net_j)' = 0.5(1-h_j^2) \tag{68}$$

From figure 4 shows that $\tau_R(k)$ is linked to the exciting nodes, VH_{j1} and $\tau_L(k)$ is linked to the exciting nodes VH_{j2} then can be calculated Jacobian vectors.

$$\frac{\partial x_m(t+1)}{\partial \tau_1'(t)} = \sum_{j=1}^{nh} W_{1j} f(net_j)' VH_{j1} \quad (69)$$

$$\frac{\partial y_m(t+1)}{\partial \tau_1'(t)} = \sum_{j=1}^{nh} W_{2j} f(net_j)' VH_{j1} \quad (70)$$

$$\frac{\partial \theta_m(t+1)}{\partial \tau_1'(t)} = \sum_{j=1}^{nh} W_{3j} f(net_j)' VH_{j1} \quad (71)$$

$$\frac{\partial x_m(t+1)}{\partial \tau_2'(t)} = \sum_{j=1}^{nh} W_{1j} f(net_j)' VH_{j2} \quad (72)$$

$$\frac{\partial y_m(t+1)}{\partial \tau_2'(t)} = \sum_{j=1}^{nh} W_{2j} f(net_j)' VH_{j2} \quad (73)$$

$$\frac{\partial \theta_m(t+1)}{\partial \tau_2'(t)} = \sum_{j=1}^{nh} W_{3j} f(net_j)' VH_{j2} \quad (74)$$

Therefore, recursive methods for calculating the Jacobian vectors are developed so that the algorithm can be applied to real-time systems. After completing the procedure from n=1 to N the new control actions for the next sample will be:

$$\tau_R(k+1) = \tau_{ref1}(k+1) + \tau_1'^K(t+N) \quad (75)$$

$$\tau_L(k+1) = \tau_{ref2}(k+1) + \tau_2'^K(t+N) \quad (76)$$

where $\tau_1'^k(t+N) \,\&\, \tau_2'^k(t+N)$ are the final values of the feedback-controlling signals calculated by the optimization algorithm. This is calculated at each sample time k so that $\tau_R(k+1) \,\&\, \tau_L(k+1)$ are torque control actions of the right and the left wheels respectively. These actions will be applied to the mobile robot system and the position and orientation identifier model at the next sampling time. The application of this procedure will continue at the next sampling time (k+1) until the error between the desired input and the actual output becomes lower than a pre-specified value.

IV. Simulation Results

The proposed controller is verified by means of computer simulation using MATLAB/SIMULINK. The kinematics and dynamics model of the nonholonomic mobile robot described in section 2 are used. The simulation is carried out by tracking a desired position (x, y) and orientation angle (θ) with a lemniscates and square trajectories in the tracking control of the robot. The parameter values of the robot model are taken from [18]: M=0.65kg, I=0.36kgm^2, L=0.105 m and r=0.033 m.

A hybrid excitation signal has been used for the robot model. Figure 7 shows the input signals $\tau_R(k)$ and $\tau_L(k)$, right and left wheel torques respectively. The training set is generated by feeding a PRBS signals, with sampling time of 0.5 second, to the model and measuring its corresponding outputs, position x and y and orientation θ.

The proposed controller is implemented based on the structure shown in figure 3. The fist stage of operation is to set the position and orientation neural network identifier. This task is performed using series-parallel and parallel identification technique configuration with modified Elman recurrent neural networks model. The identification scheme of the nonlinear MIMO mobile robot system are needed to input-output training data pattern to provide enough information about the kinematics and dynamics mobile robot model to be modelled. This can be achieved by injecting a sufficiently rich input signal to excite all process modes of interest while also ensuring that the training patterns adequately covers the specified operating region. Back propagation learning algorithm is used with the modified Elman recurrent neural network of the structure (5-6-6-3). The number of nodes in the input, hidden, context and output layers are 5, 6, 6 and 3 respectively as shown in figure 4

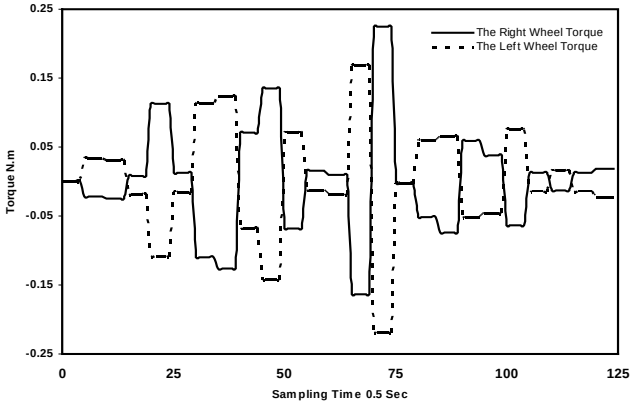


Figure 7. The PRBS input torque signals used to excite the mobile robot model.

A training set of 125 patterns has been used with a learning rate of 0.1. After 3244 epochs, the identifier outputs of the neural network, position x, y and orientation θ, are approximated to the actual outputs of the model trajectory as shown in figure 8.

Parallel configuration is used to guarantee the similarity between the outputs of the neural network identifier and the actual outputs of the mobile robot model trajectory. At 3538 the same training set patterns has been achieved with a mean square error less than 5.7×10^{-6}. The neural network identifier position and orientation outputs and the mobile robot model trajectory are shown in figure 9.

A. Case Study-1

The desired lemniscates trajectory which has explicitly continuous gradient with rotation radius changes, this trajectory can be described by the following equations:

$$x_r(t) = 0.75 + 0.75 \times \sin(\frac{2\pi t}{50}) \quad (77)$$

$$y_r(t) = \sin(\frac{4\pi t}{50}) \quad (78)$$

$$\theta_r(t) = 2\tan^{-1}(\frac{\Delta y_r(t)}{\sqrt{(\Delta x_r(t))^2 - (\Delta y_r(t))^2} + \Delta x_r(t)}) \quad (79)$$

The second stage of the proposed controller is feedforward neural controller. It uses multi-layer perceptron

neural network (8-11-2) as shown in figure 5. The trajectory has been learned by the feedforward neural controller with off-line and on-line adaptation stages using back propagation algorithm as shown in figure 6 to find the suitable reference torque control action at steady state.

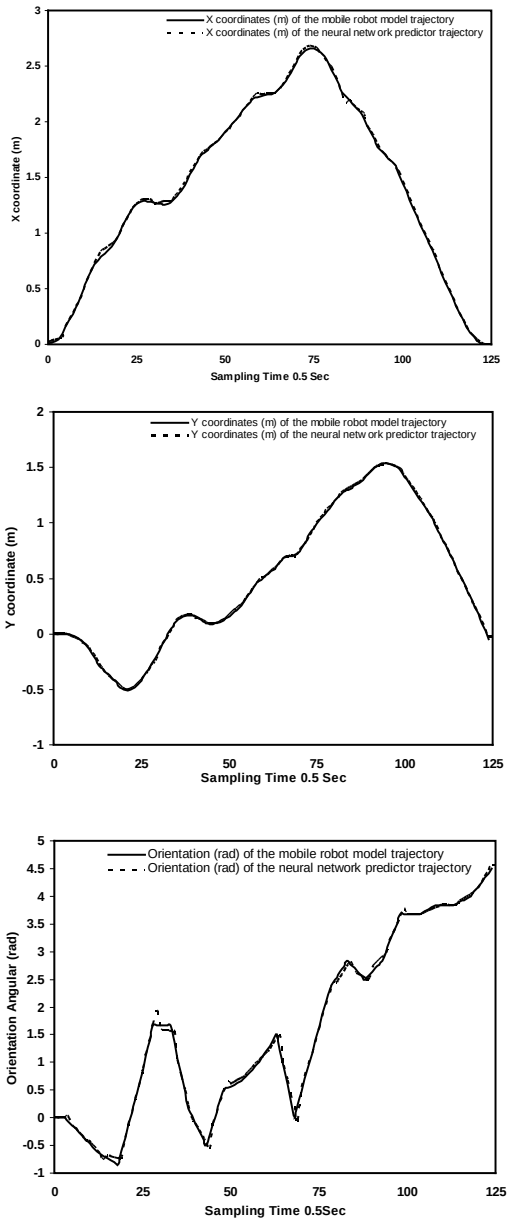


Figure 8. The response of the identifier with the actual mobile robot model output: (a) in the X-coordinate; (b) in the Y-coordinate; and (c) in the θ -orientation.

Finally the case of tracking a lemniscates trajectory for robot model, as shows in figure 3, is demonstrated with optimization algorithm for N-step-ahead prediction. For simulation purposes, the desired trajectory is chosen as described in equations 77 and 78 and the desired orientation angle is taken as expressed in equation 79. The robot model starts from the initial posture $q(0) = [0.75, -0.25, \pi/2]$ as its initial conditions.

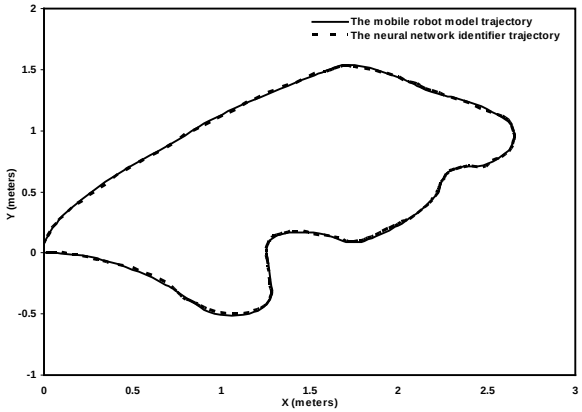


Figure 9. The response of posture identifier with the actual mobile robot model outputs for the training patterns.

A disturbance term $\overline{\pi d} = [0.01\sin(2t) \quad 0.01\sin(2t)]^T$ [5, 6 and 8] is added to the robot system as unmodelled kinematics and dynamics disturbances in order to prove the adaptation and robustness ability of the proposed controller. The feedback neural controller seems to require more tuning effort of its two parameters (Q and R). Q is the sensitivity weighting matrix to the corresponding error between the desired trajectory and identifier trajectory, while the weighting matrix R defines the energy of the input torque signals of right and left wheels. Investigating the feedback control performance of the neural predictive controller can easily obtained by changing the ratio of the weighting matrices (Q and R) as show in figure 10. This also gives the designer the possibility of obtaining more optimized control action depending on the MSE of the position and orientation, which is more difficult to obtain in other controllers. Therefore, the best value of Q parameter is equal to 0.01 and best value of R parameter is equal to 1 for obtaining more optimized control action as shown in figure 10.

The robot trajectory tracking obtained by the proposed adaptive neural predictive controller is shown in figures 11a and 11b.These figures demonstrate excellent position and orientation tracking performance for five steps ahead prediction in comparison with one step ahead prediction.
In spite of the existence of bounded disturbances the adaptive learning and robustness of neural controller with optimization algorithm show small effect of these disturbances.

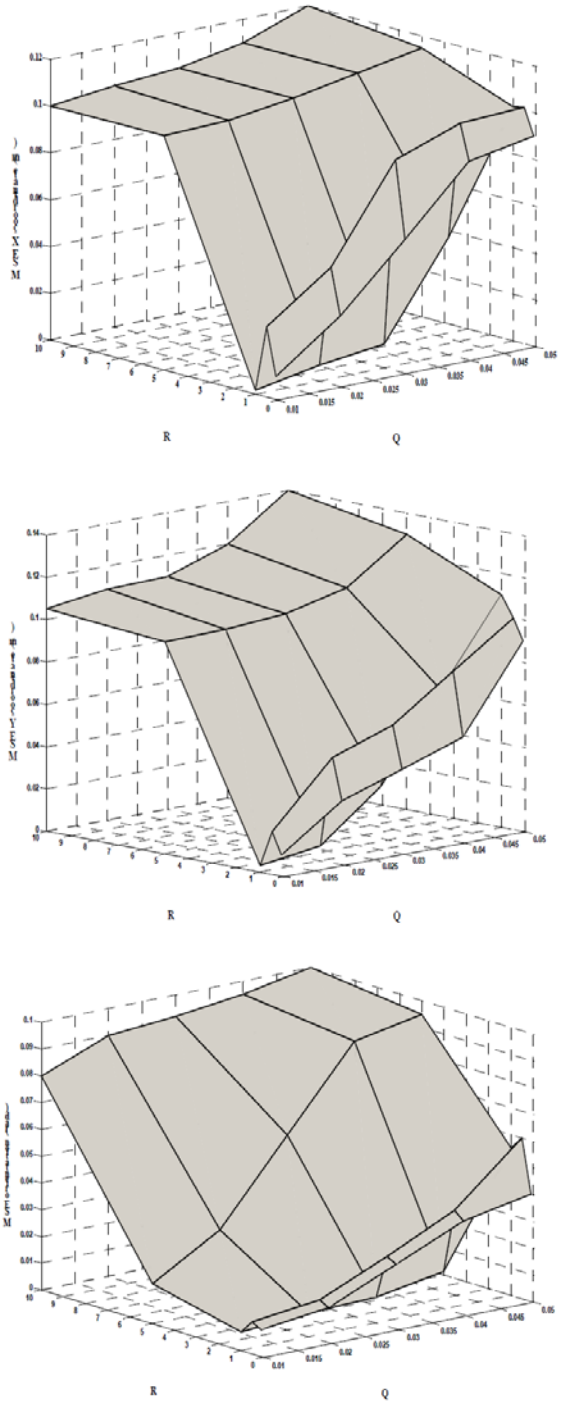
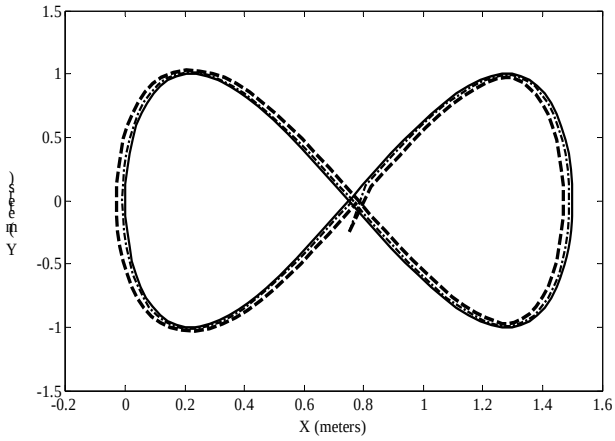


Figure 10. The MSE of position and orientation with (Q&R) parameters.



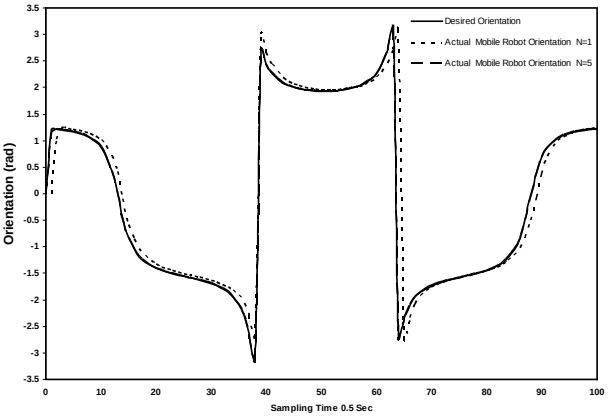


Figure 11. Simulation results for one and five steps ahead predictive: (a) actual and desired lemniscates trajectory; and (b) actual and desired orientation.

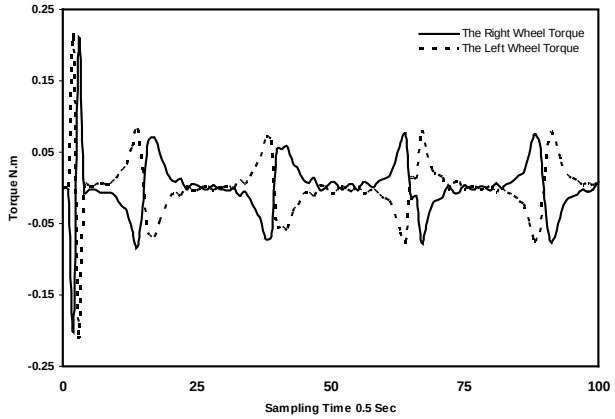


Figure 12a. The torque of the right and left wheel action for N=5.

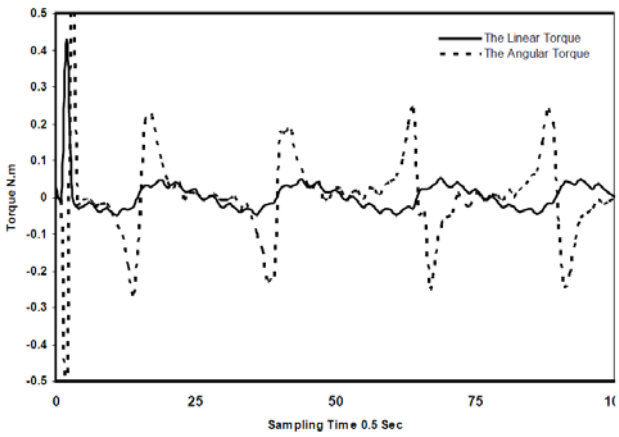


Figure 12b. The linear and angular torque action for N=5.

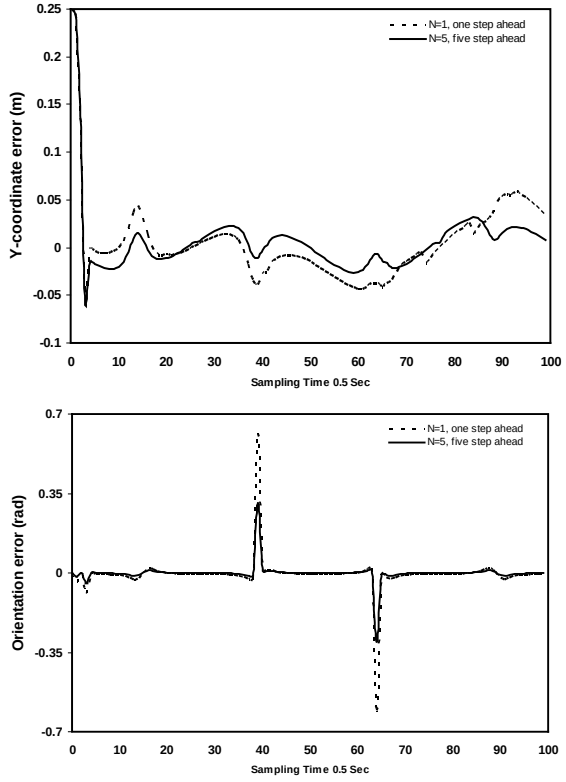


Figure 13. Position and orientation tracking error for two cases N=1, 5.

The simulation results demonstrated the effectiveness of the proposed controller by showing its ability to generate small smooth valves of the control input torques for right and left wheels without sharp spikes. The actions described in figures 12a and 12b show that smaller power is required to drive the DC motors of the mobile robot model.

The effectiveness of the proposed adaptive neural predictive control with predictive optimization algorithm is clear by showing the convergence of the pose trajectory error for the robot model motion for N=1 and 5 steps ahead as shown in figure 13.

The maximum tracking error in the X-coordinate trajectory is equal to ± 0.05m for one-step ahead while for the five steps ahead the X- coordinate error is equal to ± 0.01m. For Y-coordinate tracking error is equal to ± 0.05m for one-step ahead and for the five steps ahead the error has declined to less than 0.01m. The maximum tracking error in the orientation of the trajectory is equal to ± 0.67 radian for one-step ahead but it is equal to ± 0.34 radian for five steps ahead.

The mean-square error for each component of the state error $(q_r - q) = (e_x, e_y, e_\theta)$, for the five step ahead predictive control is $MSE(q_r - q) = (0.0012, 0.0017, 0.0387)$, while for one step ahead predictive control is $MSE(q_r - q) = (0.0021, 0.0028, 0.0577)$.

B. *Case Study-2*

Simulation is also carried out for desired square trajectory which has explicitly non-continuous gradient for verification the capability of the proposed controller performance. The mobile robot model starts from the initial position and orientation $q(0) = [0, -0.1, 0]$ as its initial posture with the same external disturbance are used in case 1 and case 2, and used the same stages of the proposed controller.

Figure 14a shows that the mobile robot tracks the square desired trajectory quite accurately but at the end of one side of the square, there is a sudden increase in position errors of the mobile robot against the desired trajectory at the corners of the square because the desired orientation angle changes suddenly at each corner as shown in figure 14b, therefore, the mobile robot takes a slow turn.

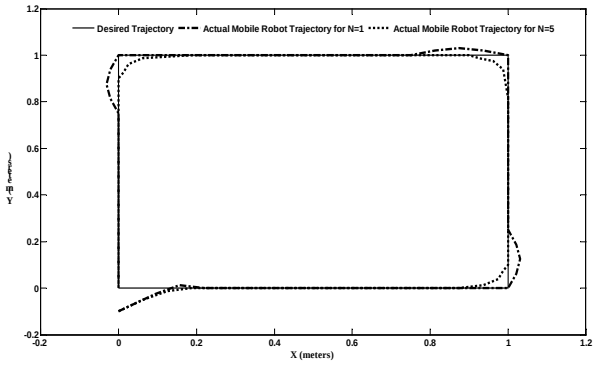


Figure 14a. Actual trajectory of mobile robot and desired trajectory for five steps ahead predictive.

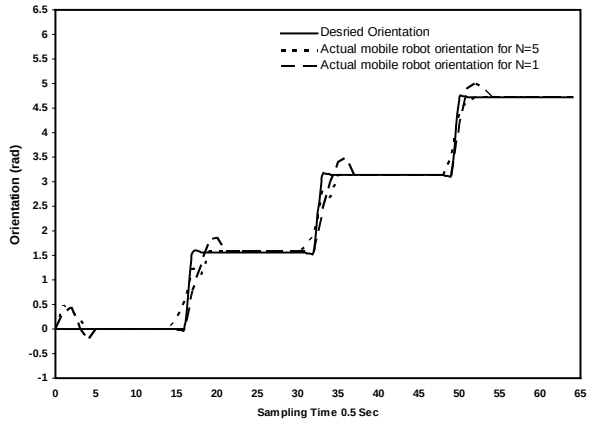


Figure 14b. Actual orientation of mobile robot and desired orientation for one and five steps ahead predictive.

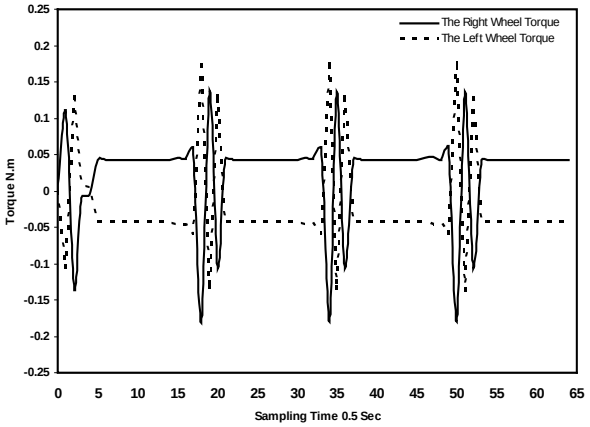


Figure 15a. The right and left wheels torque action for N=5.

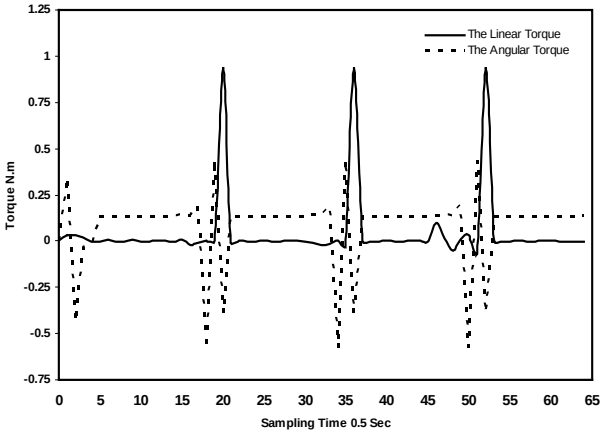


Figure 15b. The linear and angular torque action for N=5.

In figures 15a and 15b, the behaviour of the control action torques for right and left wheels is smooth values with small sharp spikes, when the desired orientation angle changes suddenly at each corner.

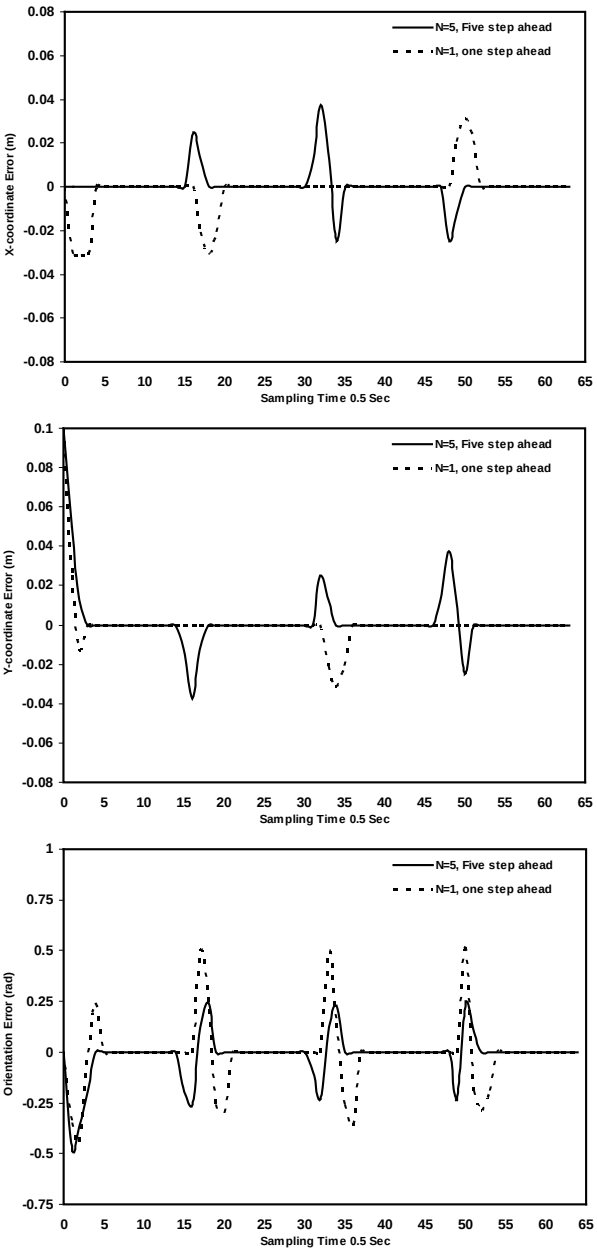


Figure 16. Position and orientation tracking error for N= 5.

In addition, the robot tracks the right side of the square desired trajectory and the tracking errors sharply drop to small values as shown in figure 16.

The maximum X coordinate error in the square trajectory is equal to ±0.03m for one step ahead predication while for the five steps ahead prediction the maximum error in the X-coordinate is equal to ±0.02m. The maximum Y coordinate error in the square trajectory is equal to -0.04m for one step ahead predication while for the five steps ahead prediction the maximum error in the Y-coordinate is equal to ±0.03m.

Along any one side of the square, the desired orientation angle is constant, therefore the orientation error is equal to zero, but at the end of one side of the square trajectory, the desired orientation angle changes suddenly, therefore, the position and orientation errors of the mobile robot against the desired trajectory at the corners of the square are increasing as shown in figure 16.

The mean-square error for each component of the state error $(q_r - q) = (e_x, e_y, e_\theta)$, for five step ahead predictive control is $MSE(q_r - q) = (0.0007, 0.0018, 0.027)$.

While for one step ahead predictive control is $MSE(q_r - q) = (0.0013, 0.0020, 0.0367)$.

From the simulation results, the five step ahead predictive gives better control results, which is expected because of the more complex control structure, and taking into account future values of the desired, not only the current value, as with one step ahead.

The main advantage of the presented approach is the analytically derived control law which has significantly high computational accuracy with optimization technique to obtain the optimal control action and to minimise tracking error of the continuous and non-continuous gradients (lemniscates and square) trajectories respectively.

V. CONCLUSIONS

The adaptive neural predictive trajectory tracking control methodology for nonholonomic wheeled mobile robot is presented in this paper. The proposed controller consists of three parts: position and orientation neural network identifier, feedforward neural controller and feedback neural controller with optimization algorithm for N-step-ahead prediction. The proposed control scheme minimises the quadratic cost function consisting of tracking errors as well as control effort. It uses two models of neural networks in the structure of the controller, multi-layer perceptron and modified Elman neural network. They are trained off-line and adapted on-line using back propagation algorithm with series-parallel and parallel configurations. Simulation results illustrated evidently that the proposed adaptive neural predictive controller model has the capability of generating smooth and suitable torque commands, τ_R and τ_L without sharp spikes. The proposed controller has demonstrated the capability of tracking continuous and non-continuous gradients desired trajectories and minimises the tracking error approximately ±0.01m for five steps-ahead prediction.

This was demonstrated when bounded external disturbances were added and achieved due to its adaptation ability and robustness behaviour.

REFERENCES

[1] R-J Wai and C-M Liu, "Design of dynamic petri recurrent fuzzy neural network and its application to path-tracking control of nonholonomic mobile robot", IEEE Transactions on Industrial Electronics, 56(7), pp. 2667-2683, 2009.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[2] F. Mnif and F. Touati, "An adaptive control scheme for nonholonomic mobile robot with parametric uncertainty", International Journal of Advanced Robotic Systems, 2(1), pp 59-63, 2005.

[3] G. Klancar and I. Skrjanc, "Tracking error model-based predictive control for mobile robots in real time", Robotics and Autonomous System, 55, pp. 460-469, 2007.

[4] F.N. Martins, W.C. Celesta, R. Carelli, M. Sarcinelli–Filho, and T.F. Bastos–Filho, "An adaptive dynamic controller for autonomous mobile robot trajectory tracking", Control Engineering Practice. 16, pp. 1354-1363, 2008

[5] T. Das, I.N. Kar and S. Chaudhury, "Simple neuron-based adaptive controller for a nonholonomic mobile robot including actuator dynamics", Neurocomputing. 69, pp. 2140-2151, 2006.

[6] J. Ye, "Adaptive control of nonlinear PID-based analogue neural network for a nonholonomic mobile robot", Neurocompting. 71, pp. 1561-1565, 2008.

[7] N.A. Martins, D.W. Bertol, E.R. De Pieri, E.B. Castelan and M.M. Dias, "Neural dynamic control of a nonholonomic mobile robot incorporating the actuator dynamics", CIMCA 2008, IAWTIC 2008, and ISA 2008, IEEE Computer Society 2008.

[8] A. S. Al-Araji, M. F. Abbod and H. S. Al-Raweshidy, "Design of a neural predictive controller for nonholonomic mobile robot based on posture identifier", Proceedings of the IASTED International Conference Intelligent Systems and Control (ISC 2011). Cambridge, United Kingdom, July 11 - 13, 2011. pp. 198-207.

[9] M.K. Bugeja, S.G. Fabri and L. Camilleri, "Dual adaptive dynamic control of mobile robots using neural networks", IEEE Transactions on Systems, Man, and Cybernetics-Part B: CYBERNETICS, 39(1), pp.129-141, 2009.

[10] Jun Ye, "Tracking control for nonholonomic mobile robots: integrating the analog neural network into the backstepping technique", Neurocomputing. 71, pp. 3373-3378, 2008.

[11] H-X Zhang, G-J Dai and H Zeng, "A trajectory tracking control method for nonholonomic mobile robot", Proceedings of the 2007 International Conference on Wavelet Analysis and Pattern Recognition Beijing, China, 2-4 Nov. 2007.

[12] S. Sun, "Designing approach on trajectory tracking control of mobile robot", Robotics and Computer Integrated Manufacturing. 21, pp. 81-85, 2005.

[13] R. Siegwart and I. R. Nourbakhah, Introduction to Autonomous Mobile Robots (MIT Press 2004).

[14] S. X. Yang, H. Yang, Q. Max and H. Meng, "Neural dynamics based full-state tracking control of a mobile robot", IEEE International Conference on Robotics & Automation, New Orleans, LA, April 2004.

[15] J. Xu, M. Zhang and J. Zhang, "Kinematic model identification of autonomous mobile robot using dynamical recurrent neural networks", IEEE International Conference on Mechatronics & Automation, Niagara Falls, Canada, July 2005.

[16] J. M. Zurada, Introduction to Artificial Neural Systems (Jaico Publishing House, Pws Pub Co. 1992).

[17] K. Ogata, Modern Control Engineering (4th Edition, by Addison-Wesley Publishing Company, Inc. 2003).

[18] K-H. Su, Y-Y. Chen and S-F. Su "Design of neural-fuzzy-based controller for two autonomously driven wheeled robot", Neurocompting. 73, pp. 2478-2488, 2010.

Q-Learning Based Traffic Optimization in Management of Signal Timing Plan

Yit Kwong Chin, Nurmin Bolong, Aroland Kiring, Soo Siang Yang, Kenneth Tze Kin Teo
Modelling, Simulation and Computing Laboratory,
School of Engineering and Information Technology,
Universiti Malaysia Sabah, Kota Kinabalu, Malaysia.
msclab@ums.edu.my, ktkteo@ieee.org

***Abstract* - Occurrences of traffic congestions within the urban traffic network are increasing in a rapid rate due to the rising traffic demands of the outnumbered vehicles on road. The effectiveness of management from traffic signal timing planner is the key solution to solve the traffic congestions, but unfortunately the current traffic light signal system is not fully optimized based on the dynamic traffic conditions on the road. Adaptable traffic signal timing plan system with ability to learn from their past experiences is needed to overcome the dynamic changes of the urban traffic network. The ability of Q-learning to prospect gains from future actions and obtain rewards from its past experiences allows Q-learning to improve its decisions for the best possible actions. A good valuable performance has been shown by the proposed learning algorithm that able to improve the traffic signal timing plan for the dynamic traffic flows within a traffic network.**

Keywords - Q-learning, signal timing plan, traffic control, learning algorithm

I. INTRODUCTION

The demands of the traffic flows in the urban area are usually supported by a complicated traffic network which covers the whole urban cities. Unfortunately, in the urban area, there are always high traffic demands with dynamic traffic conditions. In addition, reconstructions of the traffic road to cope with the high traffic demands are not an option for a fully developed city with limited availability of landscapes. When the existing traffic network is unable to meet the saturated traffic demands by the on-road vehicles, traffic congestions occurred around the urban traffic network. The most common solution for the traffic congestions problem is the implementation of traffic lights system to control the traffic flow within the traffic network.

The breakdown of the conventional traffic light~~s~~ system begins when it fails to fulfill the rising traffic demands of the traffic network. Therefore, a breakthrough evolution of the traffic lights system is needed to learn and adapt towards the dynamic characteristic of the traffic flow. The conventional method of predetermine the traffic lights signals timing plan based on the historical traffic statistics data is insufficient to handle the actual traffic flow demands.

The ability of the Q-learning system to learn from its past experiences is focused in the study of traffic signal timing plan management system. The experiences learnt by Q-learning control algorithm from its past actions will assist the algorithm to make better decisions in future for its adaption into the dynamic changes of the traffic flow within the urban traffic networks.

II. REVIEW OF TRAFFIC SIGNAL TIMING PLAN

Traffic lights systems with different traffic signal planning optimization approaches are being widely used to control the traffic flow nowadays. The system functions by coordinating the signal timing plan to ensure every phases of the traffic flow has the permission to pass through the intersection and preventing the intersection from crashed down. The paralysis of the entire traffic network can be caused by a failure at a single intersection.

Red, green and amber signals are the 3 basic signals in the traffic signal timing plans which signalize the stoppage of vehicles at intersection, permission to pass through intersection, and warning for slow down before the intersection respectively. After all the 3 signals have been given to a link in the intersection, a phase is considered completed. A cycle of traffic signal is completed~~,~~ after the traffic signal timing plan has circulated all the phases at the intersections.

Various studies have been carried out throughout these years for the enhancement of the traffic light systems through the management of the traffic signal plan. Fixed-time traffic light system is one of the primitive approaches in traffic signal timing plan, where the duration of each traffic signal is determined beforehand. The setup of the traffic signal timing system is based on the historical statistic of the traffic condition. However, this method will not allow the traffic signal timing plan to react towards the dynamic changes of the traffic condition in the variable environment of traffic networks. Therefore, artificial intelligence techniques consists of learning ability has been proposed in the evolution of the traffic light systems. Different researchers have chosen variant types of artificial

intelligence methods for the optimization of the traffic flow. Genetic algorithm or evolutionary algorithm is one of the common methods introduced into the traffic control systems. Consideration of routing of traffic flow using genetic algorithm has shown some improvement in the traffic control [1]. Fuzzy logic control is also being implemented into the traffic light systems for better control of traffic flow [2]. Enhancement of the performance of traffic light system is done with idea such as extending green light period while detecting continuously incoming vehicle flow [3]. Another approach to improve the traffic control is using wireless communications between vehicles and traffic control systems to gather information for traffic flow optimization [4]. Reinforcement learning is applied in certain studies for the traffic flow control and optimizations in recent years to model and learn the traffic behavior [5, 6, 7]. The proposed Q-learning in this study is one of the reinforcement learning algorithms that are widely used in various fields.

III. Q-LEARNING ALGORITHM

A. Basics of Q-learning

Reinforcement Learning is an algorithm that is able to improve itself through the learning process of the past. Q-Learning is one of the common methods available in reinforcement learning. In Q-learning, the exploratory agent explores in a complex and non-deterministic environment via trial-and-error approach, and then executes (exploitation) the best action based on the experience [8, 9]. Rewards or penalties will be gained by the previous trial-and error actions and stored in the memories as experiences for future references. Q-learning algorithm promised to improve the performance of an agent with the experience gained in the past.

B. Structure of Q-learning

Q-table is the main component of the Q-learning algorithm. The Q-table is a matrix table where various information extracted from the traffic plan is being stored. Each element in the Q-table is identified as Q-values, which represents a value for every single states and actions pair. The algorithm evaluates Q-value of the Q-table by (1).

The agents of Q-learning will receive a reward or penalty for every action a taken in state s with the evaluation from (1). In every iteration steps, the maximum Q-value at state s will be selected by the Q-learning agent (exploitation). Then, evaluation from the reward function for that action will be stored in the Q-table when the algorithm moves to the next state [10].

$$Q(s,a)_i = (1-\alpha)Q(s,a)_{i-1} + \alpha[R(s,a)_i + \gamma \max_{a'} Q(s',a')] \quad (1)$$

where, s = current state
a = action taken in current state
s' = next state
a' = action taken in next state
i = iteration
α = learning rate
γ = discounting factor

C. ε -Greedy Selection

Actions of the Q-learning algorithm are determined by their value of rewards or the nature of randomness. Selection by the value of rewards is a simple task of locating the maximum rewards available from the action lists. However, the selection process should be conducted in random. The random selection can be used to prevent the selection process being trapped at local maximum or minimum point.

This is where the greedy probability, ε is introduced into the algorithm. Greedy probability, ε allows the Q-learning to have a chance of randomly choose an action from the actions space which does not consist the highest Q-value [11]. The introduction of the greedy probability ensures the agent having a chance to explore into the new environment (exploration). Without the greedy probability, ε, Q-learning process will lack the tendency of exploration in the new environment which will end up with missing the chance to get in touch with a new experience with larger rewards. However, if the greedy probability, ε is too large, the stability of Q-learning will decrease as the exploration of the algorithm will hinder itself from settling down in the known environment with experience.

D. Learning Rate and Discount Rate

Learning rate and discount factor are the two elements that influence the convergence speed of the algorithm. The learning rate and discounting factor is ranged in between 0 and 1. Learning rate of the Q-learning algorithm determines the importance of the newly acquired experience. When the agent acquires a higher learning rate, where it is near to 1, then new knowledge from the learning process will have more influence towards the agent than its previous experience. If the learning rate is set to lower value or 0, then the Q-learning will not acknowledge the newly gained experience, and act upon solely on the past gained experience. This will inhibit the Q-learning algorithm from learning the environment. Thus, a suitable learning rate shall be determined for the agent to be able to learn faster in the new environment.

The discounting factor is capable of showing the degree of importance of the next state. Higher discounting factor when the value is near to 1 means the future gain prospected by the agent is more important. It might lead to faster convergence of Q-table and thus influence the overall performance of the agent. However, a high discounting factor will influence the Q-learning from

looking at the importance of the current experience. Therefore, discounting factor has to be optimum to let Q-learning algorithm has a balance focus on both the short-term rewards from the current experience and also the long-term benefit from the future prospect gain.

E. Flow Chart of Q-learning

The flow chat of the Q-learning algorithm is shown in Fig. 1. The Q-learning algorithm will start with identifying the current state s from the input. After the identification of the state s , an action will be chosen from the action list, either by searching for the maximum rewards or by random if the greedy probability ε is triggered. With all the values initialized in the previous steps, the Q-value for the action a taken in state s is calculated using (1). Q-table will then store the Q-value. In other words, the experience of the QL agent is defined inside the Q-table. The rewards and penalties of the proposed Q-Learning are evaluated by a set of simple rules of reward functions. Following the evaluating reward and updating the Q-table, the next state s' for the Q-learning algorithm will be determined after the selected of the action a is executed. Then the stopping criteria for the Q-learning algorithm will be checked with the determined next state, s'. If the next state s' is the final goal of the Q-learning, then the process will be ended, else the next state s' will become the current state s for another new iteration. The process will be continued until the goal or stopping criteria are fulfilled.

F. States and Actions

The environment model of the Q-learning algorithm is defined from the definition of the states and actions. A proper defined states and actions are crucial for a Q-learning system to ensure the exploration process can be successfully implemented throughout all the possible states. The level of queue length at each traffic phases in the intersection is chosen as the Q-learning algorithm states. The states of the Q-learning are categorized into 4 levels of queue length in this study from no queue length to high queue length. Thus, there are total of 256 possible states combination from the permutation combination of 4 phases at the intersections with 4 levels of queue length each.

The environment of Q-learning is modelled by states. In order to accomplish the exploration within the environment, execution acts that in charge of the exploration are needed. The acts are defined as actions and responsible for leading the algorithm from the current state to the other states. In this study, the green signal distributions of 1 second and 5 seconds are the action of the proposed Q-learning algorithm. The green signals are chosen by the Q-learning algorithm, and then being stored in the memory and distributed to the traffic phases after the Q-learning algorithm reaches its goal. Penalties will be given when both the actions appear to be the wrong decisions and hence no zero value of green signals is defined. The algorithm will proceed to explore the suitable green signals distribution of others traffic phases when there are certain penalties received. The penalties and rewards of each available action are evaluated in the reward and penalty functions.

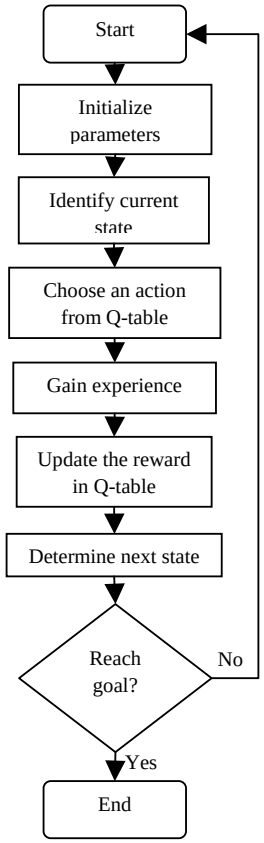


Figure 1. Q-Learning algorithm flow chart.

G. Rewards and Penalties Function

In Q-learning algorithm, each of the actions will return a certain rewards, meaning that the best action will acquire the highest rewards in the process. In order to let Q-learning decides the most optimum action without errors, proper rules or policies to reward and penalize the actions have to be carefully carried out. In traffic flow control and optimization system, the best outcome is when the system is able to archive the lowest vehicles in queue at the intersection at anytime. Thus, Q-learning algorithm must have the ability to decide the best traffic signal timing plan to produce the least number of vehicles in queue at the intersection. Appropriate rewards and penalties functions will be formed to ensure the best optimum timing plan can be built. In the proposed Q-learning algorithm, the actions are rewarded when green signal is distributed for the vehicles in the queue at the intersection. If the action is allocating unused green signal to the phase without any

vehicles waiting, a penalty will be given to the action. Theoretically, it is a waste for a longer green signal to be distributed to the phase with lesser vehicles.

During the oversaturated traffic condition, continuously incoming vehicles at a heavy traffic intersection will lengthen the green signal duration. Therefore, the Q-learning algorithm will tend to distribute longer green signal to the traffic phase, as there will always be vehicles at the intersection due to the heavy traffic flow to clear the particular traffic phase. But that decision will cause more vehicles to accumulate at the other traffic phases because of the long waiting time, and this will lengthen the vehicle queue. As a result, the action will be fined or penalized when too much green signal is allocated upon a single traffic phase. The purpose of this second penalty is to compensate and optimize the average waiting time of all the traffic users at the intersection during the saturated traffic condition. The penalty works by introducing a penalty factor into the algorithm, where the penalty factor of each traffic phase will increase for every distributed green signal. The penalty factor will become significantly large to act as a warning about the traffic phase getting too much green signals as time goes on.

However, a stopping criterion has to be set in the algorithm to indicate the accomplishment of objectives in the Q-learning algorithm. Q-learning algorithm will stop when all the traffic phases are distributed with optimum traffic signal timing plan as well as no more queue length at the intersection.

The reward results from the evaluation of the actions are continuously updated into the Q-table as Q-value for the purpose of exploration and exploitation in the future.

IV. SIMULATONS

A. Traffic Intersection

A 4-way intersection in front of University Malaysia Sabah (UMS) is used as the model in this study, which consists of 4 phases. Phases are the sequence of the traffic signals to allow only certain traffic flows to pass through the intersection at a particular time in the traffic signal timing plan management [12]. Fig. 2 shows the 4-way intersection which is labeled with 4 phases, namely phase A, B, C and D respectively. The efficiency of the traffic scheduling or traffic signals timing plan is ensured by the traffic lights setting of all the phases. Besides that, prevention of vehicle crashes at the intersections is carried out with the proper setting of the phase sequence.

B. Description of Traffic Intersection

The 4-way intersection is chosen as the simulation platform with the collected traffic data at the study site. Performance of the developed QL traffic signal timing plan management system (QLTSTM) is tested using the data collected. The results of the simulated QLTSTM system are shown in Fig. 3.

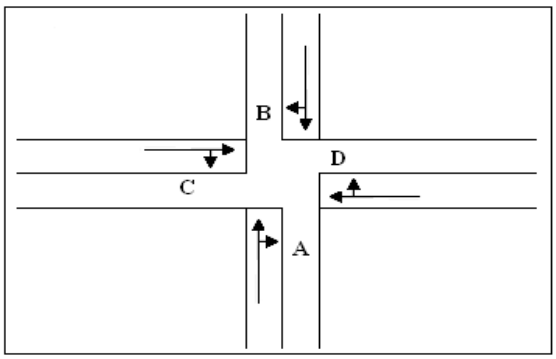


Figure 2. 4-way traffic intersection with 4 phases

Fig. 3 shows the simulation results of QLTSTM at UMS intersection. The developed QLTSTM system selected simulation time of 600 seconds to carry out the analysis for its performance. It can be observed that due to the heavy incoming vehicles at the main road, phase C and D experiencing more traffic flow than traffic phase A and B. Simulation results show that all the traffic phases manage to maintain a low level of vehicle queue length. It is proven that QLTSTM system is able to determine a suitable timing plan for the intersection.

During the traffic phase undergoing a red signal, the vehicles in queue will experience waiting time and start to accumulate which cause the slope of the graph to increase. Meanwhile, the decreasing part of the graphs show the green signal is activated for releasing the vehicles in queue to pass the intersection. From the graphs, observations show that each phases is undergoing their green signal at different timing, this indicates that only one traffic phase is given the green signal at a particular moment.

Three different situations should be tested with various traffic conditions in order to test the performance of the developed QLTSTM system. First, QLTSTM is tested and simulated with an increasing traffic demand. Then, the response of the QLTSTM on the decreasing traffic demand is also examined and analyzed through the simulation. Finally, a simulation of the QLTSTM system in the dynamic changes due to the traffic environment has been carried out to test the adaptability and robustness of the system. The results of the simulation are shown in the next section.

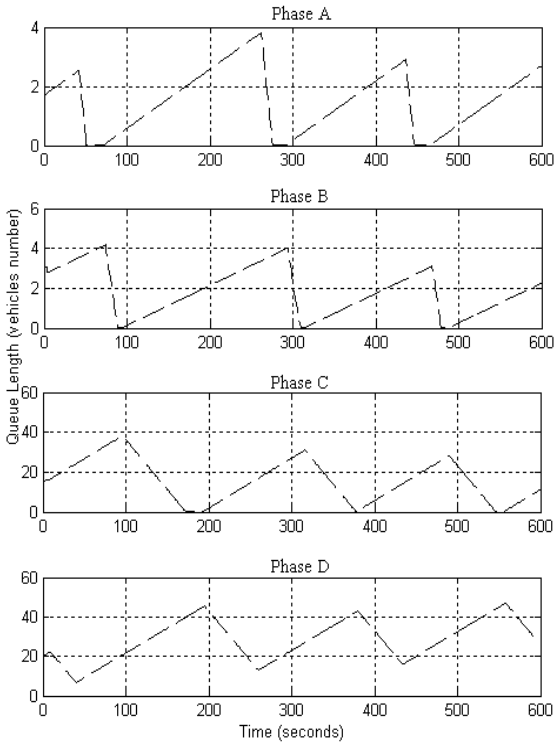


Figure 3. Simulation results of QLTSTM at UMS intersection.

V. RESULTS AND DISCUSSIONS

A. Results of Case I Simulations

Simulations of the QLTSTM have been carried out for 3600 seconds in various cases to evaluate the long term performance of the QLTSTM system. In Case I, approximately 900 seconds of an average low traffic flow is fed into the QLTSTM simulator. Then, the traffic incoming flow is started to increase until the end of the simulation. The collected practical data of the incoming traffic flow indicates the traffic condition before and during the peak hour of the day. During the first 900 seconds, the traffic condition is classified as non peak hour. The incoming flow is considered not heavy before reaching the steady state condition. After that, the incoming flow started to increase and the traffic becomes congested during the peak hour. Fig. 4 shows the simulation result of traffic phase D in Case I. Traffic phase D is chosen for analysis and discussion since it experiences the most significant changes in traffic condition at the UMS intersection. The result shows that the sudden increase of the incoming traffic flow is able to be managed by the proposed QLTSTM system. At the beginning of the simulation, QLTSTM releases most of the vehicles in queue effectively. However, it can be observed that the queue length at the traffic phase increases

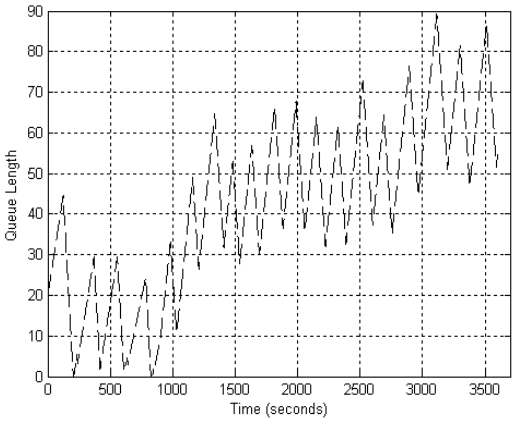


Figure 4. Simulation result of Phase D in Case I.

due to the sudden rising of the incoming flow Nevertheless, QLTSTM start to adapt in the situation by releasing more vehicles and the system manages to maintain approximately 30-60 vehicles in queue at the traffic intersection after each traffic cycle.

B. Results of Case II Simulations

Case II is referring to the traffic condition after the peak hour. The simulation begins with 900 seconds of oversaturated traffic flow and then the incoming traffic flow is gradually reduced. Fig. 5 shows the results of traffic phase D in Case II. It can be noticed that the traffic situation is oversaturated for the first 900 seconds where the maximum queue length has reached a maximum value of 65 vehicles. After the oversaturated situation, the incoming traffic flow has been reduced. The results show that the traffic condition at the intersection is eased by the QLTSTM at the later stage of the simulation, as the proposed algorithm is able to reduce the queue length at the intersection to a minimum level. At the end of every traffic phase after 1500 seconds, the vehicles in queue are reduced under 10 vehicles.

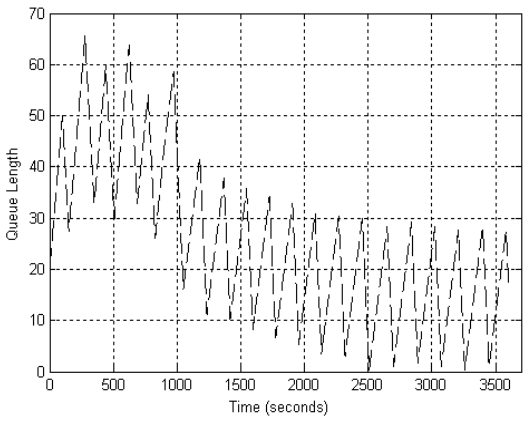


Figure 5. Simulation result of Phase D in Case II.

C. Results of Case III Simulations

Case III involves the random dynamic changes of the traffic environment and the QLTSTM is tested under this situation. Throughout the simulation, the traffic condition is varies at different traffic phases. Two heavy traffic phases are compared for a better observation on the performance of the QLTSTM system. Fig. 6 shows the results of traffic phase C and traffic phase D respectively. Both traffic phases are introduced with average traffic flow during the first 500 seconds of the simulation, and the QLTSTM performed expectedly well.

During 500 seconds to 1000 seconds, traffic phase D still maintaining the same traffic condition whereas traffic phase C is experiencing the saturated traffic flow. The QLTSTM system increases the green signal duration of traffic phase C to release more vehicles in queue as shown in Fig. 6. From time 1000 seconds to 1500 seconds, the incoming traffic reduces to average low level in both traffic phases. During the simulation time of 1500 seconds to 2500 seconds, traffic phase D has been fed with a heavy traffic flow, indicating the possibility of the begin of traffic congestion. Due to rapid increase in the incoming flow, the queue length of traffic phase D reaches 58 vehicles. However, most of the vehicles are able to be released after 3 traffic cycles via QLTSTM system's fast response towards the situation. Last part of the simulation has been carried out with the traffic phases under average low traffic flow, and the results show that QLTSTM manage to maintain the level of vehicles in queue after various traffic condition. The developed QLTSTM system's performance meets the expectation since it is able to maintain a minimum level of vehicles in queue at both traffic phases in the simulation.

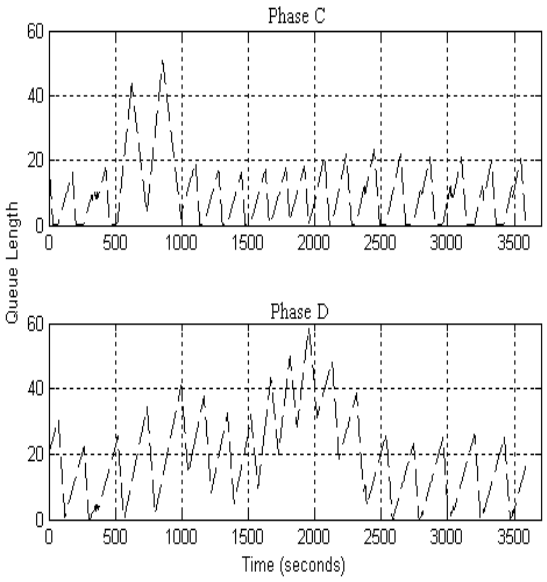


Figure 6. Simulation results of Phase C and D in Case III.

D. Discussions

During simulation of Case I where the traffic condition refer to the peak hour of the day, optimal green signal duration gained from the QLTSTM system is able to optimize and control the traffic situation. QLTSTM still release most of the vehicles waiting at the intersection, even though the queue length of the traffic phase D cannot maintain at the minimal level as before the oversaturated traffic condition.

QLTSTM system determines the green signal duration for traffic phase D in this situation to reduce the average waiting time of other traffic phases for the aim of traffic flow optimization. Based on the observation in Fig. 4, QLTSTM does not continuously to distribute green signal for traffic phase D as the result of its ability to compromise own interest. The system only allocates 50 seconds of green time for traffic phase D in each traffic cycle, even there are still vehicle in queue. The penalties in the reward function of the Q-learning algorithm restrict itself to distribute too much green signal duration to certain traffic phase. The ability of QLTSTM to defend the global benefits instead of local interest is verified throughout the simulation result.

QLTSTM system's capability to react fast towards the traffic condition is shown in the results of Case II. Although during the first 900 seconds, heavy traffic flow has burdened the traffic phases where a massive amount of vehicles is accumulated, QLTSTM system still successfully maintains its level of vehicles in queue and reacts fast to the traffic flow changes. Based on Fig. 5, 6 traffic cycles are used by QLTSTM system within 1000 seconds and 2000 seconds to release most of the vehicles in queue at the intersection right after it detects the changes of the traffic condition.

The adaptability of the system to the dynamic environment of the traffic networks is focused in the final study of the simulation. Various combinations of incoming traffic flow have been implemented to evaluate the QLTSTM system. Based on the observation of Fig. 6, investigation on the effect of heavy traffic flow towards other traffic phases has been carried out. A sudden heavy traffic flow changes applied to traffic phase C during 500 seconds to 1000 seconds where the vehicle in queue has a sudden rise over 40 vehicles. QLTSTM system has to allocate more green signals duration for traffic phase C in order to cope with the traffic demands. Vehicle queue length at traffic phase D has experienced a slight increase because of the action of QLTSTM in phase C. It is reasonable since the red signals duration of other traffic phases will increase as longer green signal is allocated to a particular traffic phase. Thus, the ability of the QLTSTM system to reward and penalize towards the action via the reward function has been evaluated and assessed.

VI. CONCLUSIONS

In this paper, studies have been carried out on the traffic flow control systems. The developed Q-learning based traffic signal timing plan management system has shown its performance through the simulations and the ability to perform well in various traffic environments are proven.

The simulation results verify the QLTSTM system has the ability for optimizing the traffic phase to utilize less green signal durations. With the purpose of reducing the average waiting time and queue length of the other traffic phases, QLTSTM compromises the green signal durations of one traffic phase for others. Fast reaction towards the changes of traffic flow input is also one of the capabilities of QLTSTM. The QLTSTM system's sensitivity towards the dynamic environment allows it to adapt in the dynamic changes of the incoming traffic flow as well as the vehicles in queue.

Q-learning algorithms' exploration in the dynamic traffic flows and exploit its best actions based on its experience has shown a good performance in the traffic signal timing plan management system. The traffic signal timing plan system is encouraged by the ability of Q-learning to learn and adapt with the dynamic changes of the traffic flow. Q-learning is assessed via the simulation to be a suitable method or technique to be implemented into the traffic flow control and optimization of urban traffic network system.

ACKNOWLEGDEMENT

The authors would like to acknowledge the financial assistance of the Ministry of Higher Education of Malaysia (MoHE) under Fundamental Research Grant Schemes (FRGS) No. FRG0105-TK-1/2007 and FRG0220-TK-1/2010, University Postgraduate Research Scholarship Scheme (PGD) by Ministry of Science, Technology and Innovation of Malaysia (MOSTI).

REFERENCE

[1] F. Teklu, A. Sumalee, and D. Watling. "A Genetic Algorithm Approach for Optimizing Traffic Control Signals Considering Routing." *Computer-Aided Civil and Infrastructure Engineering*, vol. 22, pp. 31-43, 2007.

[2] E. Azimirad, N. Pariz, and M.B.N. Sistani. "A Novel Fuzzy Model and Control of Single Intersection at Urban Traffic Network." *IEEE Systems Journal*, vol. 4, no. 1, pp. 107-111, March 2010.

[3] K. Khiang Tan, M. Khalid, and R. Yusof. "Intelligent Traffic Lights Control by Fuzzy Logic." *Malaysian Journal of Computer Science*, vol. 9, no. 2, pp. 29-35 1996.

[4] V. Gradinescu, C. Gorgorin, R. Diaconescu, and V. Cristea. "Adaptive Traffic Light using Car-to-Car Communication." *In Proceeding of Vehicular Technology Conference, 2007,* pp.21-25.

[5] I. Arel, C. Liu, T. Urbanik, and A.G. Kohls. "Reinforcement Learning-based Multi-Agent System for Network Traffic Signal Control." *IET Intelligent Transport Systems*, vol. 4, no. 2, pp. 128-135, 2010.

[6] Z.Y. Liu, and F.W. Ma, "On-line Reinforcement Learning Control for Urban Traffic Signals." *In Proceedings of the 26th Chinese Control Conference*, 2007, pp. 34 – 37.

[7] P.G. Balaji, X. German, and D. Srinivasan, "Urban Traffic Signal Control using Reinforcement Learning Agents." *IET Intelligent Transport Systems*, vol. 4, no. 3, pp. 177-188, 2010.

[8] B.Abdulhai, R. Pringle, and G.J. Karakoulas. "Reinforcement Learning for True Adaptive Traffic Signal Control." *Journal of Transportation Engineering*, vol. 129, no.3, pp. 278-285, June 2003.

[9] C.J.C.H. Watkins, P. Dayan. "Technical Note: Q-learning." *Machine Learning*, vol. 8, no.3, pp. 279-292, May 1992.

[10] Y.K. Chin, L.K. Lee, N. Bolong, S.S. Yang, and K.T.K. Teo, "Exploring Q-Learning Optimization in Traffic Signal Timing Plan Management." *In Proceeding of 3rd International Conference on Computational Intelligence, Communication Systems and Networks*, 2011, pp. 269-274.

[11] Mi. Tokic, and G. Palm. "Value-Difference based Exploration: Adaptive Control between Epsilon-Greedy and Softmax." *In KI 2011: Advances in Artificial Intelligence*. J. Bach, S.Edelkamp, Ed. Berlin: Springer, 2011, pp. 335-346.

[12] N.J. Garber, and L.A. Hoel. *Traffic and Highway Engineering. 3rd Ed.* Pacific Grove, California: Thomson, 2002.

Adapting MCMC with CUSUM Path Plot for Overlapped Vehicle Tracking

Wei Yeang Kow, Wei Leong Khong, Hoe Tung Yew,
Ismail Saad, Kenneth Tze Kin Teo
Modelling, Simulation and Computing Laboratory,
School of Engineering and Information Technology,
Universiti Malaysia Sabah, Kota Kinabalu, Malaysia.
msclab@ums.edu.my, ktkteo@ieee.org

***Abstract* - Traffic surveillance using video sensors has been essential over the recent years and it is capable of obtaining wide range of vehicle information. However tracking overlapped vehicles still remain as a challenging task due to the involvement of high dimensional calculation. Conventional fixed sample size Markov Chain Monte Carlo (MCMC) will encounter tracking error if the sample size is insufficient and will be computationally expensive if the sample size is too large. Therefore cumulative sum (CUSUM) path plot is introduced to aid the difficulties in determining the sample size of MCMC. The adaptive sample size of MCMC has shown significant tracking accuracy especially when the vehicle is overlapped. Furthermore, implementation of observation likelihood by fusing colour and edge distance has further enhanced the tracking performances. Experimental result shows that CUSUM path plot algorithm has overcome the limitation of fixed sample size MCMC with better tracking accuracy and lesser computational time.**

Keywords - Markov Chain Monte Carlo, cumulative sum, vehicle tracking, adaptive sampling

I. INTRODUCTION

The capability of video sensor in obtaining wide range of vehicle information has become essential nowadays and it is widely implemented on vehicle tracking for traffic flow control purposes [1]. This sensor has the capability of obtaining information such as vehicle's motion, size, colour and different types of observable parameters which are applicable for vehicle tracking purposes. As a result, close-circuit television system (CCTV) has been implemented to capture the traffic flow and the traffic behaviour is monitored via the traffic light management system. Various strategies have been carried out in vehicle tracking system to determine the traffic trajectories for applications on sophisticated traffic surveillance [2]. It uses modelled, feature, region and active contour based strategies to track vehicles and has shown promising tracking results. However, tracking the overlapping vehicle still remains as a challenging issue since it requires high dimensional calculation and it is difficult to be solved. MCMC is one of the widely implemented algorithms in object tracking. The sampling efficiency of MCMC is critical in predicting the vehicle position since appropriate sample size is difficult to be determined.

This paper outline consists of seven sections. Section 2 is the literature review of various MCMC approaches in vehicle tracking. Section 3 describes the methodology of the MCMC tracking algorithm. Section 4 discusses the computation of CUSUM path plot in diagnosing the convergence of MCMC and Section 5 is the implementation of CUSUM path plot in MCMC to track the overlapped vehicle. Results and discussions are made at Section 6 and the last section presents the conclusion of this paper.

II. LITERATURE REVIEW

Various studies have been performed on tracking vehicle with MCMC particle filtering algorithm and the implementation results show it is capable of tracking the vehicle in different occluded conditions. However, the performances consist tracking error due to its inappropriate foreground-background segmentation method [3]. Implementation in [4] has shown that detection of overlapped vehicle using shape and width of predefined vehicle model is significant and accurate. Nevertheless, the algorithm will encounter confusion if the vehicle is occluded with another vehicle that has the same size and contour feature. Data Association MCMC is well known on its capability in tracking multiple objects using six moves which are birth, split, death, merge, reduce and extend movement [5, 6]. The tracking performance is accurate but the algorithm is time consuming and will have failure when the vehicle is overlapped for a long period. The research in [7] has developed MCMC tracking algorithm that can track overlapped vehicle from front view without the background segmentation and shadow elimination method. It uses simulated annealing to adaptively sample the MCMC and has successfully tracked the overlapped vehicle. However the developed algorithm can only perform tracking on vehicle that view from the top front which their background is the non complicated traffic lane. It will encounter tracking error if the vehicle is segmented at a complicated background especially from the side view where there are trees and other environmental obstacle that affect the tracking accuracy.

Other than those mentioned approaches, most of the MCMC tracking algorithm are based on fixed sample size which are inefficient since MCMC tracking performance

are highly dependent on its sampling efficiency [5, 6, 8]. Small sample size will leads to tracking error due to insufficient information whereas large sample size provides better tracking accuracy with compensation of higher computational time. Therefore appropriate sample size are require to track overlapped vehicle efficiently where higher sample size will be developed when the vehicle is overlapped while smaller sample size will be used if the vehicle are not overlapped and easier to be tracked.

In this paper, CUSUM path plot is proposed to adaptively sample the MCMC for tracking overlapping vehicle. The algorithm is capable of quantitatively determine the convergence rate of MCMC by computing the hairiness of the plot [9, 10]. Furthermore, observation likelihood combining both colour and edge distance enable the algorithm to detect and track overlapped vehicle that has high similarity on the shape and outlook. Fixed sample size MCMC with CUSUM path plot based adaptive MCMC are implemented to track overlapped vehicle and their corresponding performances are discussed and analyzed.

III. MCMC

MCMC is a Bayesian inference computation that is capable of solving high dimensional problem [11]. The algorithm estimates the vehicle state and calculates its prior probability $P(\theta)$, proposal distribution $Q(\theta)$, and observation likelihood $\pi(\theta)$. The new estimated state sample θ^* will be proposed based on the previous accepted state sample θ_t^{i-1} where i is the MCMC sample index and t is the frame index. The estimated sample will be determined to be accepted into the MCMC based on the Metropolis-Hasting acceptance ratio, α as shown in (1).

$$\alpha = \min\left(1, \frac{p(\theta^*)Q(\theta_t^{i-1} \mid \theta^*)\pi(\theta^*)}{P(\theta_{t-1}^i)Q(\theta^* \mid \theta_t^{i-1})\pi(\theta_t^{i-1})}\right) \quad (1)$$

The proposed sample is accepted with the probability $\alpha = 1$ or else the sample is rejected and the previous accepted sample will be re-generated and accepted as new sample in MCMC. The sample is being estimated and proposed until it reached the stopping criteria that have been set. The developed MCMC will then evaluate its mean value as shown in (2). The variable n is the number of samples in the MCMC which is also defines as the MCMC sample size. θ_t^i is the accepted sample of the MCMC and the expected mean in (2) is the estimated vehicle state which is also indicating the tracked vehicle position.

$$E[\theta] = \frac{1}{n}\sum_{i=1}^{n}\left(\theta_t^i\right) \quad (2)$$

A. Prior Probability Distribution

Prior probability distribution calculates the probability of acceptance of the current estimated vehicle state sample based on the finalized state of previous video frame. The vehicle state space is $\theta = \{x, y\}$ where x and y is the centroid coordinates of the vehicle. Lower probability value will be computed if the estimated sample is far from the distribution range and less likely to be accepted into MCMC. The prior probability is shown in (3).

$$p(\theta) = \frac{1}{2\pi\sigma_p^{\,2}} e^{-\frac{(\theta-\theta_{t-1})^2}{2\sigma_p^{\,2}}} \quad (3)$$

Equation (3) is a 2-dimensional Gaussian distribution and θ_{t-1}^i is the mean of vehicle state at the previous tracking frame. Gaussian distribution is used due to its characteristic in providing balance distribution range and is able to determine the vehicle state that will be estimated at unpredictable direction. $\sigma_p^{\,2}$ is the variance of the distribution which will determine the range of state sample that can be proposed. Larger estimation range indicates that the variance is set to large value and smaller range indicates smaller variance value.

B. Proposal Distribution

Proposal distribution is being used to estimate the new state sample. It is a 2-dimensional Gaussian distribution and the new sample is estimated based on the previous accepted sample. The proposal distribution is defined in (4).

$$Q(\theta \mid \theta_t^{i-1}) = \frac{1}{2\pi\sigma_q^{\,2}} e^{-\frac{(\theta-\theta_t^{i-1})^2}{2\sigma_q^{\,2}}} \quad (4)$$

Variable θ_t^{i-1} is the previous accepted vehicle state sample. The distribution will compute lower probability value if the estimated sample is apart from the previous accepted sample. This is necessary to prevent the algorithm from estimating the sample that is out of the desired tracking range. Metropolis-Hasting acceptance ratio will calculate the acceptance probability of the proposed sample along with the observation likelihood and prior probability. If the proposed sample is accepted, it will be used as the reference state sample to compute another proposed state sample for the next iteration.

C. Observation Likelihood

Observation likelihood is the measurement of similarity of the proposed vehicle state sample to the target vehicle model. Edge and color features are common parameters that have been widely implemented in machine vision to determine the outlook characteristics of objects [12]. Therefore the implemented likelihood is based on the edge distance transform and RGB color similarity [7, 13]. The actual vehicle model is extracted and it is used as the reference to calculate the observation likelihood. If the proposed state is near to the target vehicle then the similarity is higher. Higher likelihood value is produced and MCMC will be more probable to accept the proposed sample. Fig. 1 shows the color and edge vehicle model that have been extracted.

RGB color histogram has been constructed to compute the color likelihood. The histogram is built with bin size of $8\times8\times8$ where the bin combination will increase the likelihood sensitivity. This has enabled the accurate comparison of colour histogram between the proposed sample and the model. Bhattacharyya coefficient is used as the indicator to determine the colour histogram similarity which is defined in (5).

$$\rho(p,q) = \int p(u)q(u)du \tag{5}$$

The bin index is indicated by u . $q(u)$ is the constructed colour histogram of vehicle model and $p(u)$ is histogram of the proposed vehicle state. Bhattacharyya distance will then be calculated using (6).

$$d = \sqrt{1-\rho(p,q)} \tag{6}$$

Large Bhattacharyya distance value indicates that the proposed sample is similar to the target vehicle model and low distance value signifies the sample is not similar to the target vehicle. The computation of colour likelihood is defined in (7). If the proposed sample has high similarity or small Bhattacharyya distance value, the colour likelihood value will be increased and the proposed sample is more probable to be accepted into MCMC.

$$\pi(C\,|\,\theta) = \left(\sqrt{1-\rho(p,q)}\right)^{-1} \tag{7}$$

For the calculation of edge likelihood, edge distance transform has been implemented. Edge distance transform calculates the distance of the pixels from the vehicle edge. Thus the distance value will be larger if the proposed vehicle state is further apart from the actual vehicle's edge. The steps of computing the edge distance using the vehicle

(a) (b)

Figure 1. (a) Colour model (b) Edge distance model

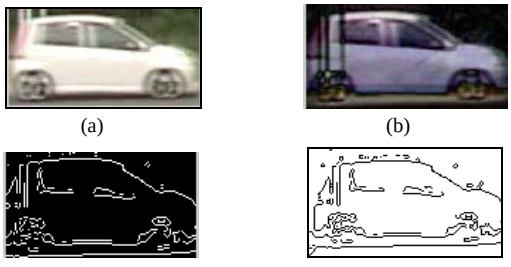
model is shown in Fig. 2. Fig. 2(a) is the extracted target vehicle model. The absolute difference between the background and vehicle model are illustrated in Fig. 2(b). The vehicle edge is then extracted using Sobel edge detection algorithm which is shown in Fig 2(c). Finally the edge distance transform is performed on the extracted edge by computing the Euclidean distance to the nearest neighbour edges as displayed in Fig. 2(d). The edge likelihood is then determined with (8) and (9).

$$d = \frac{1}{n}\sum_{n=1}^{T} EdgeDist(x,y) \tag{8}$$

$$\pi(E\,|\,\theta) = \frac{1}{2\pi\sigma_d^{\,2}} e^{-\frac{d}{2\sigma_d^{\,2}}} \tag{9}$$

From (8), the variable (x,y) is the coordinate of the edge pixels of the vehicle model. T is the total number of edge pixels and the corresponding coordinate of edge distance of the proposed sample is added up to calculate the distance average. Edge distance likelihood is determined using (9) which is a Gaussian distribution function. Hence if the proposed sample is near to the target vehicle, the distance d is small and the edge distance likelihood value will be increased.

Fusion of colour and edge distance has been implemented to compute the observation likelihood. The combinations of both parameters enable the algorithm to keep track on the overlapped vehicle. When vehicle is fully overlapped at the back of another vehicle, the algorithm can only track the front vehicle based on the edge distance likelihood. After the overlapped vehicle begins to appear, colour likelihood assists to keep tracking on the target



(a) (b)

(c) (d)

Figure 2. Edge Distance Transform

vehicle. Thus weighting between these two likelihoods is essential in the tracking algorithm. The defined observation likelihood is shown in (10) where β_θ and γ_θ is the weight of colour and edge distance likelihood respectively. The colour and edge distance likelihood priority can be set by calibrating the weight to suits the overlapping situations.

$$\pi(\theta) = \beta_\theta\left[\pi(C \mid \theta)\right] \cdot \gamma_\theta\left[\pi(E \mid \theta)\right] \tag{10}$$

D. MCMC Tracking Algorithm

The vehicle tracking algorithm is implemented on the image frame of the video data. An image is extracted from the video to track the target vehicle and the first vehicle state θ_t^1 will be initialized. By using the proposal distribution, a new vehicle position θ^* is proposed where computation on prior probability $P(\theta)$ and observation likelihood $\pi(\theta)$ are performed. If the proposed position is near to the target vehicle and met the computed Metropolis-Hasting acceptance ratio then the state will be accepted as a new sample θ_t^i . However the previous accepted sample θ_t^{i-1} will be replicated as the new sample in MCMC if the current proposed state did not achieve the acceptance ratio. The iterative sampling process will be repeated until the stopping criterion is met. Vehicle position is estimated by calculating the expected mean value of the MCMC using (2) and the computed position will then be used for prior probability calculation of the next tracking frame.

Moving to the next frame, the first MCMC sample is initialized by randomly select one of the samples from the previous frame. This can ensure that the initial state are near to the target vehicle and will not vary too far from the estimated position of the previous frame. New sample will be proposed again and the MCMC procedure is repeated until the entire video frames have been processed. TABLE I shows the implemented algorithm which is based on fixed sample size. The sample size is predetermined and the algorithm will stop sampling and extract the vehicle position when the sample size limit is reached.

The conventional MCMC tracking algorithm is lack of computation efficiencies as appropriate sample size are needed to be determined before the tracking is performed. Tracking with large sample size will increase computational time whereas setting sample size that is too small will cause tracking error. Thus, CUSUM path plot is implemented to adaptively determine the appropriate MCMC sample size that suits the tracking situation. The proposed algorithm is capable of enhancing the tracking performance with higher accuracy and lesser processing time.

TABLE I. MCMC TRACKING ALGORITHM

MCMC Tracking Algorithm
Initial first MCMC sample θ_t^1 at time *t* with $P(\theta)$
for frame = 1 to *t*
Initial first sample θ_t^1 at time *t* from θ_{t-1}
for *i* = 1 to *n*
Estimate new state θ^* with proposal distribution *Q*
Calculate prior probability, $P(\theta_{t-1}^i)$ and $P(\theta^*)$
Calculate observation likelihood, $\pi(\theta_t^{i-1})$ and $\pi(\theta^*)$
Compute the acceptance ratio, α
if $\alpha = 1$ **then** add $\theta_t^i = \theta^*$
else add $\theta_t^i = \theta_t^{i-1}$
end if
end for
Calculate vehicle position $E[\theta]$
end for frame

IV. CUSUM PATH PLOT

CUSUM path plot is capable of quantitatively determine the convergence rate of MCMC. The MCMC will stop sampling or stop proposing new samples when the CUSUM path plot diagnosed MCMC have been converged. The convergence rate is indicated by the hairiness of CUSUM path plot. Hairiness is the mixing rate of MCMC samples and it is calculated based on the maximum and minimum tuning point of the state samples' Euclidean distance [9, 10]. MCMC is diagnosed as converged if the rate of occurrence of turning point increases vastly and reaches CUSUM stopping criteria. The CUSUM path plot algorithm first calculates the mean of MCMC sample until the latest accepted sample θ_t^i as shown in (11). Euclidean distance of each accepted sample is then calculated based on the computed mean on every state acceptance's iteration as defined in (12).

$$\mu = \frac{1}{n}\sum_{i=1}^{n} \theta_t^i \tag{11}$$

$$S_i = \sum_{i=1}^{n} \left(\theta_t^i - \mu\right) \tag{12}$$

The hairiness of CUSUM path plot is determined by plotting the Euclidean distance S_i . If the plot is smooth

then MCMC are diagnosed as undergoing slow mixing rate and still require more sample to reach convergence. In another way, hairy plot indicates that MCMC is in fast mixing rate approximating to convergence [10, 14]. The hairiness of plot is determined by using the hairiness index, D_i as shown in (13).

$$D_i = \begin{cases} 1 & if \quad S_{i-1} > S_i \quad and \quad S_i < S_{i+1} \\ & or \quad S_{i-1} < S_i \quad and \quad S_i > S_{i+1} \\ \frac{1}{2} & if \quad S_{i-1} == S_i \quad and \quad S_i == S_{i+1} \\ 0 & otherwise \end{cases} \tag{13}$$

The hairiness index will be added by 1 if the plot encounter local minimum and local maximum turning point. If the Euclidean distance remains constant and did not undergo changes then hairiness index is incremented by 0.5. If it is undergoing a smooth plot, then increment will not be performed on the hairiness index and MCMC will not reach convergence. The hairiness is calculated as shown in (14).

$$H = k\left(\frac{1}{n}\right)\sum_{i=1}^{n} D_i \tag{14}$$

The constant k is implemented for the calibration of hairiness value which enables the algorithm adaptable to various tracking situations and dynamic environmental effects. As the MCMC sample size n grows larger, the law of large numbers states that the hairiness will approximate to normal distribution with the mean value of 0.5. Hence MCMC can be diagnosed as converged if the hairiness is within the boundary as defined in (15).

$$H = \frac{1}{2} \pm Z_{\frac{\alpha}{2}} \sqrt{k\left(\frac{1}{4n}\right)} \tag{15}$$

The boundary in (15) indicates that MCMC will reach convergence when the computed hairiness is approximated to mean $1/2$ with variance of $k(1/4n)$. Variable $Z_{\alpha/2}$ is the defined confidence interval that determines the amount of samples required within the boundary [14]. To ensure most of the accepted states are accurate, MCMC is identified as converged when 95% of the samples are within the boundary. Equations (16), (17) and (18) are used to determine $Z_{\alpha/2}$.

$$1 - \alpha = 0.95 \tag{16}$$

$$\phi\left(Z_{\frac{\alpha}{2}}\right) = 1 - \frac{\alpha}{2} = 0.975 \tag{17}$$

$$Z_{\frac{\alpha}{2}} = \phi^{-1}(0.975) = 1.96 \tag{18}$$

Equation (16) calculates the confidence measure where 95% of the MCMC samples fall within the boundary. The cumulative distribution of $Z_{\alpha/2}$ is computed in (17) and will be further evaluated as shown in (18). Equation (15) will then be updated by substituting (18) and the new boundary equation is defined in (19).

$$H = \frac{1}{2} \pm 1.96 \sqrt{k\left(\frac{1}{4n}\right)} \tag{19}$$

MCMC is diagnosed as converged when hairiness fulfils the boundary condition and then the algorithm will stop MCMC from generating new samples. The converged MCMC sample size is suitable for further evaluation since the evaluated vehicle position is more accurate compared to the fixed size MCMC.

V. CUSUM PATH PLOT TRACKING

CUSUM path plot is implemented into the conventional MCMC tracking algorithm to adaptively sample MCMC and it has the similar procedure as fixed sample size MCMC. To further enhance the tracking performance, higher-order prior probability distribution is implemented [15]. The proposed distribution is capable of overcoming the limitation of the conventional first order prior probability.

Conventional MCMC encounters tracking error when the vehicle undergoes long overlapping situation. The tracking error occurs when the occluded vehicle is fully covered by the other vehicle. When the overlapped vehicle is partially appear, the prior probability is calculated based on the previous state θ_{t-1} where the vehicle is still fully occluded. The correct information is lost during the overlapping and the prior probability might not be computed based on the correct position of the occluded vehicle. To overcome this problem, *m*-th order prior probability distribution is introduced.

The distribution calculates the prior probability based on the previous *m*-th frame, before the vehicle is overlapped. The multiple frame information will shift the prior probability to a more reliable and accurate position based on the vehicle state from the previous *m*-th frame. It enables the algorithm to track the vehicle that is undergoing long overlapping period. However the selection of the *m*-th order is essential for the tracking

performances. Selecting *m*-th order that are too large might lead to tracking error as the vehicle position might change a lot over multiple frames whereas smaller order might not be enough to overcome the information lost in the long period of overlapping. Hence the selection of the *m*-th order must be carried out according to the overlapping situation. The proposed prior probability is defined in (20) and (21).

$$\theta_h = \theta_{t-1} + \frac{1}{m}\sum_{k=1}^{m}\theta_{t-k} \tag{20}$$

$$P(\theta \mid \theta_{t-1:t-m}) = \frac{1}{2\pi\sigma_p{}^2} e^{-\frac{(\theta-\theta_h)^2}{2\sigma_p{}^2}} \tag{21}$$

Equation (20) is used to replace θ_{t-1} in (3) and develop the proposed prior probability in (21). The proposed prior probability will be implemented on both fixed sample size MCMC and CUSUM path plot MCMC and their performances will be compared. TABLE II is the algorithm of the CUSUM path plot MCMC in tracking the overlapped vehicle.

VI. RESULTS AND DISCUSSIONS

The tracking of overlapped vehicle under fixed sample size MCMC and CUSUM path plot MCMC have been performed. Fixed sample size MCMC is simulated using 30 MCMC sample size and 100 MCMC sample size and their performances are analyzed as compared to CUSUM path plot MCMC. Fig. 3 shows the tracking results of CUSUM path plot MCMC. The tracking accuracy of the algorithm is then determined by using Euclidean distance as shown in Fig. 4. Fig. 5 displays the adaptive sample size of CUSUM path plot MCMC. The Euclidean distance in Fig. 4 is the difference of the tracked vehicle position to the actual vehicle position. The vehicle is determined as lost track if the Euclidean error exceeds the value of 40 whereas error below that indicates the algorithm is still able to keep track on the vehicle.

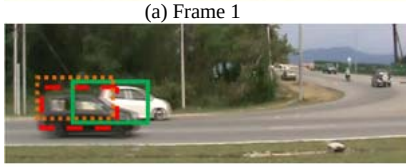
In Fig. 3, the vehicle tracked by CUSUM path plot MCMC is indicated by the line bracket. The dotted bracket and the line bracket are the tracking results of MCMC with 100 sample size and 30 sample size respectively. Every MCMC based algorithms have shown good tracking results in Frame 1 when the tracking vehicle is free from occlusion. After the vehicle is started to overlap at Frame 4, the CUSUM algorithm is still capable of tracking the vehicle compared to MCMC with 30 and 100 sample size. Both 30 and 100 sample size consist of higher tracking error since the estimated position has been influenced by another vehicle that appears to block the target tracking vehicle. When the target vehicle is occluded, the RGB colour information of the target vehicle is blocked by the occluded vehicle at front and the computed value of colour likelihood will decreased vastly. The edge distance likelihood of the vehicle appears in front of the occluded vehicle has affect the computation of observation likelihood of the proposed sample and become more probable to be accepted near to its area.

TABLE II. CUSUM PATH PLOT MCMC TRACKING ALGORITHM

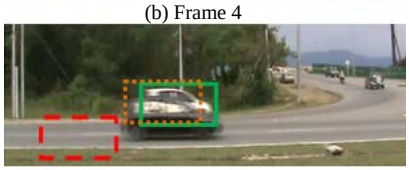
CUSUM Path Plot MCMC Tracking Algorithm

1: Initial first MCMC sample θ_t^1 at time *t* with $P(\theta)$
2: **for** frame = 1 to *t*
3: Initial first sample θ_t^1 at time *t* from θ_{t-1}
4: **Loop**
5: Estimate new state θ^* with proposal distribution *Q*
6: Calculate prior probability, $P(\theta_t^{i-1} \mid \theta_{t-1:t-m})$ and $P(\theta^* \mid \theta_{t-1:t-m})$
7: Calculate observation likelihood, $\pi(\theta_t^{i-1})$ and $\pi(\theta^*)$
8: Compute the acceptance ratio, α
9: **if** $\alpha = 1$ **then** add $\theta_t^i = \theta^*$
10: **else** add $\theta_t^i = \theta_t^{i-1}$
11: **end if**
12: Sum up current sample set $\sum_{i=1}^{n} \theta_t^{\,i}$
13: Compute μ and S_i
14 **if** found local maximum or minimum point
15: add 1 to hairiness index
16: **else if** 3 consecutive S_i remain constant
17: add ½ to hairiness index
18: **else**
19: add 0 to hairiness index
20: **end if**
21: Compute hairiness *H*
22: **if** *H* lies within $\frac{1}{2} \pm 1.96\sqrt{k\left(\frac{1}{4n}\right)}$
23: go to end Loop
24: **else**
25: go to Loop
26: **end if**
27: **end Loop**
28: Compute vehicle position $E[\theta]$
29: Record computed state θ_t for *m*-th order prior distribution calculation.
30: **end for frame**

(a) Frame 1



(b) Frame 4



(c) Frame 7

(d) Frame 12

Figure 3. MCMC CUSUM path plot vehicle tracking

Hence it can be observed in Frame 4 that fixed size MCMC tends to track the front blocking vehicle due to the decrease of colour likelihood value. Nevertheless, the capability of CUSUM path plot to keep tracking on the vehicle is due to the implementation of the third order prior probability distribution is shifted to a more accurate position according to the previous frame of the vehicle movement. The shifting of the distribution allows the proposed samples placed near to the target vehicle are more probable to be accepted into MCMC and therefore produces a better tracking position.

The target vehicle is almost completely overlapped at Frame 7 and the CUSUM algorithm is still capable of tracking the target vehicle accurately. This is due to the adaptive sampling of CUSUM path plot where appropriate MCMC sample size has been computed for the computation of vehicle position. MCMC with 30 sample size has totally lost track on the target vehicle whereas MCMC with 100 sample size has shown inaccurate tracking result since its prior probability is poorly distributed. The 100 sample size MCMC is also over sampled where more defected samples are accepted since the complete overlap situation will lead to computation of inaccurate observation likelihood, which lack of color likelihood feature and affected by the edge likelihood of the front blocking vehicle.

At Frame 12, the target vehicle has just completed the overlapping and it is tracked accurately with CUSUM path plot using small MCMC size since there are no disturbances that will affect the observation likelihood calculation. However the dotted tracker has shifted beyond the target vehicle due to the oversize of MCMC where

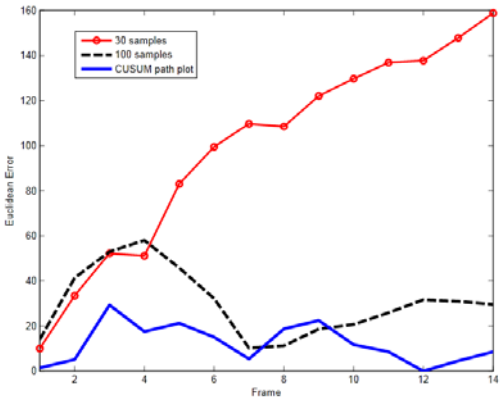


Figure 4. Euclidean distance of vehicle position

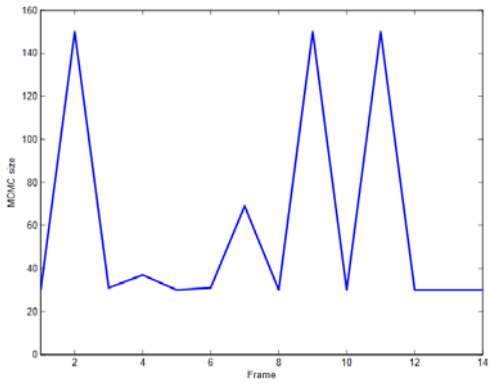


Figure 5. MCMC converged sample size

sample that are defected have been accepted and increased along with the growing sample size.

The overall tracking performance of Fig. 3 is further elaborated and represented with Euclidean distance in Fig. 4. It can be seen that MCMC with 30 sample size has the highest tracking error. The tracking encounters failure at the early frame when the vehicle is overlapped at Frame 2 and the algorithm is not able to track the vehicle since then. The main reason of the failure is due to the insufficient sample size which reduces the possibility of MCMC to accept better likelihood samples and will cause tracking error if the accepted samples are defected at the early stage. The 100 sample size MCMC represented by dash plot has shown better tracking performances compared to the 30 sample size. However, the algorithm encounters higher tracking error when the vehicle is overlapped within Frame 2 and Frame 8. The likelihood information is lost when the target vehicle is occluded and the poorly distributed prior probability has caused the tracking error. Higher sample size MCMC can be implemented to overcome this limitation but it is computationally expensive and bares the risk of increasing defected sample which is not desirable in the tracking algorithm.

Results in Fig. 4 justify CUSUM path plot has better tracking accuracy compared to the 100 sample size MCMC with lower Euclidean error. It can be observed in Fig. 5

that CUSUM path plot MCMC has adapted the MCMC to track the vehicle according to the tracking situation. When the vehicle is not overlapping, CUSUM MCMC tends to generate lesser sample size to track the vehicle. This is because small sample size is sufficient enough to calculate the exact vehicle state and the vehicle can be tracked easily. More sample size is generated when the vehicle begins to overlap and just after the overlapping. During these critical situations, the accepted sample is unstable due to the insufficient information of the target vehicle. The computation of hairiness in (14) shows that the plot is not hairy enough due to the inconsistency of the MCMC samples' mean and is not yet determined as converged where more sample states are required to improve the tracking accuracy.

For instance, the algorithm has stop sampling at the sample size of 150 at Frame 9 and 11 when the vehicle is partially overlapped. The prior vehicle information at this stage is still insufficient because of the long overlapping period. Hence there will be a lot of accepted MCMC samples that are defected and not yet suitable for vehicle position extraction. Therefore additional sample have been proposed to compensate the defected MCMC states until the hairiness has reached within the boundary in (19). Experimental results show that CUSUM path plot MCMC consumes 31% lesser computational time than 100 sample size MCMC. As a result, CUSUM path plot has met a better trade off between accuracy and efficiency of tracking performance.

VII. CONCLUSION

The CUSUM path plot MCMC algorithm produces better performance in tracking the overlapping vehicle with lesser processing time. The proposed algorithm is capable of adaptively determining the sample size in MCMC by generating more samples when the vehicle is occluded and less sample size is utilized to track the vehicle. The ability to cope with environmental disturbances by adaptively alters the sample size in CUSUM path plot has overcome the limitation of the conventional fixed size MCMC. Observation likelihood fusing both color and edge distance likelihood has been introduced and the implementation of *m*-th order prior probability has further enhanced the algorithm tracking accuracy. Thus it can be concluded that the proposed CUSUM path plot MCMC algorithm is efficient in tracking the overlapping vehicle.

ACKNOWLEDGEMENT

The authors would like to acknowledge the financial assistance of the Ministry of Higher Education of Malaysia (MoHE) under Fundamental Research Grant Schemes (FRGS) No. FRG0220-TK-1/2010, and the University Postgraduate Research Scholarship Scheme (PGD) by Ministry of Science, Technology and Innovation of Malaysia (MOSTI).

REFERENCES

[1] H.Y. Cheng, P.Y. Liu and Y.J. Lai. "Vehicle tracking in daytime and nightime traffic surveillance videos," in *Education Technology and Computer*, 2010, pp. 122-125.

[2] B. Coifman, D. Beymer, P. McLauchlan and J. Malik. "A real-time computer vision system for vehicle tracking and traffic surveillance," in *Transportation Research Part C: Emerging Technologies*, 1998, pp. 271-288.

[3] F. Bardet and T. Chateau. "MCMC particle filter for real-time visual tracking of vehicles," in *Intelligent Transportation Systems*, 2008, pp. 539-544.

[4] X. Song and R. Nevatia. "A model-based vehicle segmentation method for tracking," in *Computer Vision*, 2005, pp. 1124-1131.

[5] S. Oh, S. Russell and S. Sastry. "Markov Chain Monte Carlo data association for multiple-target tracking." *Automatic Control*, vol. 54, pp. 481-497, 2009.

[6] Q. Yu, I. Cohen, G. Medioni and B. Wu. "Boosted Markov Chain Monte Carlo data association for multiple target detection and tracking," in *Pattern Recognition*, 2006, pp. 675-678.

[7] Y. Jia and C. Zhang. "Front-view vehicle detection by Markov Chain Monte Carlo method," in *Pattern Recognition*, 2009, pp. 313-321.

[8] H.X. Xia, N.W. Yao, Z. Wei, Z. Jiang and F.Y. Xiao. " Multi-object visual tracking based on reversible jump Markov Chain Monte Carlo." *Computer Vision* , vol. 5, pp. 282-290, 2011.

[9] P. Brooks and G. Roberts. "Convergence assessment techniques for Markov Chain Monte Carlo." *Statistics and Computing*, vol. 8, pp. 319-335, 1998.

[10] P. Brooks. "Quantitative convergence diagnosis for MCMC via CUSUMS." *Statistics and Computing*, vol. 8, pp. 267-274, 1998.

[11] C. Andrieu, N. Freitas, A. Doucet and M. Jordan. "An introduction to MCMC for machine learning." *Machine Learning*, vol. 50, pp. 5-43, 2003.

[12] P.N. Trung, W.J. Kang and S.H. Ong. "Fusing color and contour in visual tracking," in *Machine Vision Application*, 2005, pp. 9-12.

[13] W.L. Khong, W.Y. Kow, F. Wong, I. Saad and K.T.K. Teo. "Enhancement of particle filter approach for vehicle tracking via adaptive resampling algorithm," in *Computational Intelligence, Communication Systems and Networks*, 2011, pp. 259-263.

[14] S.E. Adlouni, A.C. Favre and B. Bobee. "Comparison of methodologies to assess the convergence of Markov Chain Monte Carlo methods." *Computational Statistics and Data Analysis*, vol. 50, pp. 2685-2701, 2005.

[15] P. Pan and D. Schonfeld. "Visual tracking using high-order particle filtering," in *Signal Processing Letter*, 2011, pp. 51-54.

Overlapped Vehicle Tracking via Enhancement of Particle Filter with Adaptive Resampling Algorithm

Wei Leong Khong, Wei Yeang Kow, Lorita Angeline,
Ismail Saad, Kenneth Tze Kin Teo
Modelling, Simulation and Computing Laboratory,
School of Engineering and Information Technology
Universiti Malaysia Sabah, Kota Kinabalu, Malaysia.
msclab@ums.edu.my, ktkteo@ieee.org

***Abstract* - Traffic surveillance and on-road security have elevated the demand of machine vision aided traffic control system. Through the modern video camera technology, vehicle tracking has become a vital approach to assist the on-road traffic systems. In the past, many tracking methods have been developed based on the detail and information extracted from the captured vehicle. However, conventional tracking system need to be improved since the background noises and sudden appear objects will increase the difficulties of continuously tracking the target vehicle. Hence, a particle filter algorithm with adaptive resampling approach has been proposed to overcome the vehicle occlusion problems. In addition, the proposed resampling approach can also be used to solve the common particle degeneracy problem. Experimental results show that the enhanced particle filter equipped with adaptive resampling algorithm is significantly improving the accuracy of the tracking process without compromising the processing time.**

Keywords - particle filter, adaptive re-sampling, vehicle tracking, traffic surveillance

I. INTRODUCTION

Vehicle tracking is an essential approach that has drawn the attention among the researchers due to its numerous fields of applications such as road traffic control system, traffic surveillance and security system [1]. However, occlusion and overlapping between vehicles is a challenging task in surveillance system via image processing. Due to the difficulties and complexity caused by the occlusion problems, the researchers are incited to study the effective and efficient vehicle tracking method. In this research, video sensor is chosen as the tracking infrastructure rather than others sensors because of the rapid development in video camera technology. Furthermore, a wide range of the information to describe the target vehicle such as the colour, motion, edge, shape and speed of the vehicle can be extracted from the video sensor via image processing techniques.

Alternatively, vehicle flow consists of dynamic changes which could lead to non-linear and non-Gaussian conditions. Hence, particle filter has been chosen as the vehicle tracking algorithm in this research due to its ability to overcome the non-linear and non-Gaussian situations. Nevertheless, particle degeneracy was the main factor that will influence the accuracy of vehicle tracking results. Therefore, an efficient and effective resampling approach will be needed to solve the common particle degeneracy problem. Thus, an enhancement of the particle filter with the adaptive resampling step is implemented to continuously tracking the target vehicle under various overlapping incidents without compromising the processing time.

II. REVIEWS OF OBJECT TRACKING

In the past, many methods have been developed for object tracking purpose. Among the well known techniques of image processing tracking methods are Markov chain Monte Carlo [2], Kalman filter [3, 4], optical flow and particle filter [5, 6, 7]. Vehicle flow consists of dynamic changes may lead to non-linear and non-Gaussian situation. In this case, the development of extended version of Kalman filter can be used to overcome the non-linear situation. However, when the nonlinearity is inaccurately approximated by the extended Kalman filter technique, the estimated results will be diverging and hence lead to an inaccurate tracking result. Particle filter is proven as a promising and powerful technique to overcome the non-linear and non-Gaussian situation. It has been chosen as the overlapped vehicle tracking technique due to its ability to cope with the non-linear and non-Gaussian state [8, 9].

In research [10], the non-rigid objects have been tracked by using the colour feature. It is suggested that the algorithm selected for object tracking purpose should be able to deal with the partial occlusion and scale invariant incidents. Nevertheless, colour is a powerful feature that can be implemented in these situations. The extracted colour histogram of the target vehicle will be comparing with the colour histogram of the sample vehicle by using Bhattacharyya distance. As discussed in [10], the colour based algorithm can efficiently handle the non-rigid and fast moving objects under different conditions.

According to reference [11], the typical particle filter will face a phenomenon named as particle degeneracy during the tracking process. Particle filter degeneracy occurs due to the low weight or weak particle is selected

after several iterations and it blocks the further improvement of the algorithm. In general, there are two ways used in solving the particle degeneracy problem. The solution is either increasing the number of particle size implemented in the algorithm or resampling the particles. However, increase the particle size is insufficient due to the huge amount of the sample size could lead to higher computational complexity and processing time. On the other hand, resampling can increase the accuracy of the tracking results by eliminating the low weight particles and regenerate with strong weight particles without compromising the processing time [12].

As mentioned in [13], occlusion and sudden appear objects incident are among the challenging tasks that will be faced in vehicle tracking. Therefore, the proposed particle filter algorithm must be robust whether in the partially or fully occlusion incidents. When the target object is being overlapped, the information of the target object will be vanished or influenced by the obstacle. The tracking performance based on the results is accurate but the algorithm fails when there is a large degree of disjoint during the tracking.

III. METHODOLOGY

This section discusses the methodology of the particle filter and explains the likelihood of the samples being computed.

A. Particle Filter Framework

In this section, a brief review of particle filter will be presented. Particle filters also known as sequential Monte Carlo which is an iterative process that estimates posterior distribution from a finite set of weighted particles [14]. It is developed based on estimating the current state of the target vehicle from the previous particle set. Basically, the conventional particle filter algorithm consists of three important steps which are prediction stage, measurement stage and resampling stage.

In the prediction stage, the particle filter will generate a new particles set with each particle represents the estimated posterior position. The increment of the number of particles can improve the accuracy of the estimation. Meanwhile, the computational time will be longer with the large amount of the particle size.

In the measurement stage, each particle weight is computed based on likelihood probability. For visual tracking purpose, the observation state of the target vehicle can be colour, edge or shape that extracted from the information of the target vehicle. In this study, colour feature has been selected as the parameter for the vehicle tracking.

The third stage refers to the resampling process. Resampling is an important step to reduce the particle degeneracy problems. Particle degeneracy affects the accuracy of the tracking results. During particle degeneracy, the low weight particle is continually selected and used by the particle filter algorithm which leads to inaccuracy results. Hence, resampling algorithm is needed to eliminate the low weight particle and regenerate a new set of particles until the particle sets with large weight is obtained. It is important to avoid the particle degeneracy problem to improve the accuracy of the tracking results.

The dynamic changes in vehicle tracking usually consist of nonlinear and non-Gaussian elements. The posterior probability density function as described in (1) can be obtained through the prediction stage. However, the observation probability density function as in (2) is used to express the likelihood of the colour feature.

$$p(X_t \mid Z_t) \tag{1}$$

$$p(Z_t \mid X_t) \tag{2}$$

The quantities of the tracked object is denotes by state vector X_t while all the observations state at time t is denotes by vector Z_t. In the prediction stage, the prior probability density function can be obtained through (3) and the posterior probability density function is defined using the Bayes' rule in (4).

$$p(X_t \mid Z_{1:t-1}) = \int p(X_t \mid X_{t-1}) p(X_{t-1} \mid Z_{1:t-1}) dX_{t-1} \tag{3}$$

$$p(X_t \mid Z_{1:t}) = \frac{p(Z_t \mid X_t) P(X_t \mid Z_{1:t-1})}{p(Z_t \mid Z_{1:t-1})} \tag{4}$$

In the particle filter algorithm, the posterior probability density function developed from the prior density is represented by a set of N weighted particle samples. Meanwhile, posterior density function can be obtained through (5) because the weighted particles are in discrete nature and w_t^i is the normalized weight as shown in (6).

$$p(X_t \mid Z_{1:t}) \approx \sum_{i=1}^{N} w_t^i \delta(X_t - X_t(i)) \tag{5}$$

$$w_t^i = w_{t-1}^i \frac{p(z_t \mid x_t^i) p(x_t^i \mid x_{t-1}^i)}{q(x_t \mid x_{1:t-1}, z_{1:t})} \tag{6}$$

B. Color Distribution Model

In this research, colour feature has been chosen for the vehicle tracking purpose because of its ability to deal with the partial occlusion and scale invariant problems. Moreover, the processing time to obtain the colour information of the target vehicle is much faster than other

parameters. Hence, colour feature has been chosen and implemented in most of the visual tracker.

The colour histogram of the target vehicle is normally calculated in the RGB colour space to obtain a discrete 8 × 8 × 8 bins histogram. After obtaining the colour histogram of the target vehicle, it will be compared with the colour histogram of the reference vehicle in order to compute the similarity of the two histograms which is called likelihood. Bhattacharyya distance is a common technique used to measure the likelihood between two colour histograms. After the obtaining the Bhattacharyya distance, the weight of the particle can be calculated based on the likelihood.

C. Bhattacharyya Distance

Generally, the measurement between two colour distribution histogram is calculated by the common technique called Bhattacharyya coefficient [15]. In this case, the Bhattacharyya coefficient is used to calculate the coefficient between colour histogram of reference vehicle, $p = \{p_u\}_{u=1\ldots N_c}$ and colour histogram of the target vehicle, $q = \{q_u\}_{u=1\ldots N_c}$ as shown in (7).

$$\rho[p,q] = \int \sqrt{p_u q_u}\, du \tag{7}$$

Since the colour histogram is a discrete density model, the Bhattacharyya coefficient can be obtained through (8).

$$\rho[p,q] = \sum_{u=1}^{N_c} \sqrt{p_u q_u} \tag{8}$$

The value of Bhattacharyya coefficient represents the similarities of two colour distribution. The larger value of the coefficient means more similarity in the colour distribution. However, the limit of the coefficient is set from 0 to 1. If both the histogram is identical, then the coefficient will be indicated as 1.

After obtaining the Bhattacharyya coefficient, the Bhattacharyya distance can be computed using (9).

$$b_{dist} = \sqrt{1 - \rho[p,q]} \tag{9}$$

Based on the Bhattacharyya distance, the weight of the particles can be calculated using (10) where σ is the adjustable standard deviation which can be chosen experimentally.

$$\varphi_c = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{b_{dist}^2}{2\sigma^2}} \tag{10}$$

The weight of the particles is set to heavy when the colour of the reference vehicle and the colour of the target vehicle are too similar. This estimated position represented by the heavy particle will become the possible location of the target vehicle and it will be updated in the measurement stage of the particle filter algorithm.

IV. RESAMPLING ALGORITHM

As discussed in the previous section, resampling process is needed to eliminate the particle degeneracy problem. This section will describe the particle degeneracy in more detail. Then the conventional resampling and proposed resampling algorithms are discussed towards the particle degeneracy situation.

A. Particle Degeneracy

Particle filter is a good approach in vehicle tracking due to the ability to deal with non-linear and non-Gaussian situations. However, after several iterations, the particle filters also facing the problem caused by particle degeneracy. Generally, implementing the particle filter algorithm with large size of particle samples or resampling the particle samples is capable of avoiding the particle degeneracy problem. Although both approaches can be used to improve the accuracy of the vehicle tracking, resampling is more suitable to be implemented into the particle filter algorithm. Resampling is chosen since the computational time is much lesser compared to applying huge number of particles.

In order to measure the appearance of the particle degeneracy problem, the effective sample sizes need to be calculated using (11). Since, the true weight of particles in (12) cannot be determined, an estimate of the effective sample size will be computed using (13).

$$N_{eff} = \frac{N_s}{1 + Var(w_t^{*i})} \tag{11}$$

where

$$w_t^{*i} = \frac{p(x_t^i \mid z_{1:t})}{q(x_t^i \mid x_{t-1}^i, z_t)} \tag{12}$$

$$\hat{N}_{eff} = \frac{1}{\sum_{i=1}^{N_s} (w_t^i)^2} \tag{13}$$

The w_t^i in (13) is calculated using (6) and it is the normalized weight of each particle. After obtaining the estimated effective sample size, (13) is used to indicate whether particle degeneracy problem occurs or not.

If $\hat{N}_{eff} < N_s$, the particle degeneracy problem has occurred and resampling is needed. The particle filter will keep on recursively resampling until the requirement is reached.

B. Conventional Resampling

Resampling is commonly used to overcome the particle degeneracy problem. In the past, various type of resampling algorithm has been developed. The most common resampling algorithms being used to reduce the particle degeneracy are multinomial resampling, residue resampling, systematic resampling and stratified resampling.

In Table I, $\hat{N}_{eff}$ is the estimated effective sample size which is used to determine whether the particle degeneracy occurs or not. When the particle degeneracy occurs during the vehicle tracking process, the estimated effective sample size will be less than the initiated sample size. As a result, resampling will be activated and a new set of particle samples is regenerated. The newly generated particle samples will be reweighted via the colour likelihood technique. If the new set of the particle samples is not able to reach the threshold of the effective sample size, the resampling process is then continually being evaluated until the requirement is fulfilled.

When the overlapping vehicle is tracked by using the traditional resampling algorithm, the tracking results will not be promising. This is because when the target vehicle is being occluded by the obstacle, the information that can be extracted by particle filter will be limited. Therefore, more iteration of resampling algorithm is needed to obtain an accurate tracking result. However, after a few iteration of resampling process, the information of the vehicle will be influenced by the obstacle and hence the target vehicle will be lost track. Besides that, more computation time is consumed if the particle filter is recursively repeating the resampling stage due to the dissimilarity colour histogram of the reference vehicle and the colour histogram of the target vehicle.

C. Proposed Resampling

In this study, an enhanced particle filter algorithm with the adaptive function of resampling is proposed. With the proposed resampling algorithm, the resampling computational time will be optimized in vehicle tracking under various overlapping conditions. The enhanced resampling approach only resamples the low weight particles where the weight is below the threshold value. Meanwhile, the particles with the accepted weight will be stored as the reference position for vehicle tracking purpose. This is to shorten the resampling process and hence reduce the computational time with a promising vehicle tracking results. In short, it is suitable to be used when the vehicle is without occlusion or overlapping.

TABLE I. CONVENTIONAL RESAMPLING ALGORITHM

PF with Conventional Resampling Algorithm
1: Initialize reference colour histogram and sample size
2: FOR FRAME = 1, 2,…,*t*
3: **PREDICTION:**
4: FOR i = 1, 2,…, N_s
5: Draw predicted particles from prior dynamics
6: Compute the colour histogram based on estimated position
7: END FOR
8: **MEASUREMENT:**
9: Calculate the Bhattacharyya distance, b_{dist}
10: Compute the weight of the particle based on Bhattacharyya distance, φ_c
11: Normalize the weight, $w_t^i = w_t^i (\sum_{i=1}^{N} w_t^i)^{-1}$
12: Calculate $\hat{N}_{eff}$
13: WHILE $\hat{N}_{eff} < N_s$
14: **RESAMPLING:**
15: Resample the discrete distribution
16: Generate new set of particles
17: Compute the weight of the particle based on Bhattacharyya distance, φ_c
18: Check the effective sample size
19: END WHILE
20: **OUTPUT:**
21: Obtain the position of the target vehicle
22: **END FOR FRAME**

However, when the target vehicle is being occluded, the information extracted by the particle filter algorithm might be incorrect and hence it might wrongly track the target vehicle due to lack of resampling process. Thus, during the occlusion, only the largest weight of the particle will be reserved for resampling purpose. Based on the largest weight of the particle, it is fast to track the target vehicle even only a small portion of the vehicle's color is visible.

Moreover, the proposed resampling algorithm can also assist in gaining back the information of the target vehicle faster than other resampling algorithm. The enhanced resampling algorithm for vehicle tracking is shown in Table II. Meanwhile, the flowchart of the particle filter with the adaptive resampling algorithm is illustrated in Fig. 1.

V. TARGET LOCALIZATION

In this section, estimation of the location for the target vehicle will be discussed. After the entire particles are weighted by the likelihood, the particle filter algorithm will

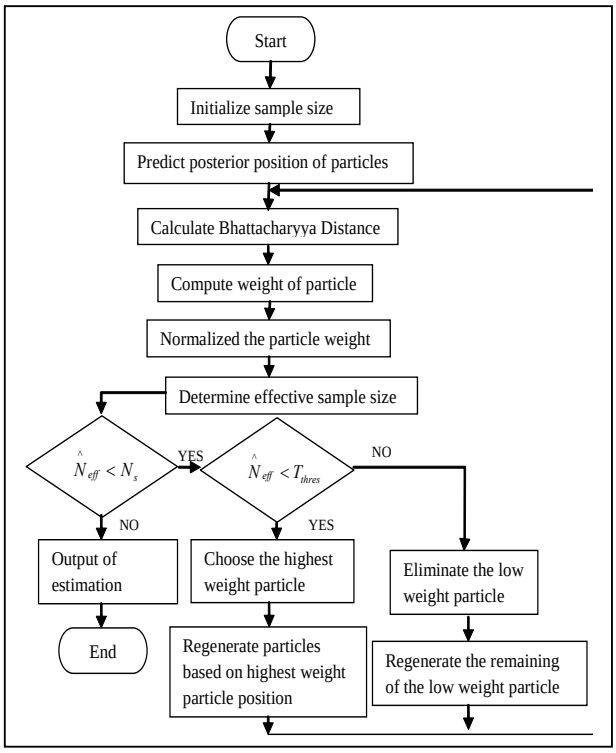


Figure 1. Flowchart of particle filter with adaptive resampling algorithm.

determine the needs of the resampling step. If resampling step is executed, the algorithm will check for the occlusion incident via the threshold. When the estimated effective sample size is lower than the threshold, the target vehicle is partially occluded or fully occluded. Hence, the highest weight of the particle will be remained as the reference position for the next prediction step while eliminating others particles. This action can speed up the algorithm to gain back the information of the target vehicle when there are a small portion colour of the target vehicle is visible.

On the other hand, when the estimated effective sample size is higher value than the threshold, the target vehicle is not occluded. Thus, in the resampling step, the particle filter only eliminate the low weight particle and at the same time the heavy weight particle will be remained. After that, the number of eliminated particle will be resampled and this step is repeating until the estimated effective sample size fulfils the requirements. The enhanced resampling process shortens the computational time compared to the conventional resampling algorithm because it only resamples the particles which potentially cause the degeneracy problem.

After resampling, output stage exists as the estimated position of the target vehicle. The position of the target vehicle can be easily calculated through the mean of all the coordination generated by the particles. This is because after the resampling stage, all the particles will be focused and concentrated at the centre of the target vehicle.

TABLE II. ADAPTIVE RESAMPLING ALGORITHM

PF with Adaptive Resampling Algorithm

1: Initialize reference colour histogram and sample size
2: FOR FRAME = 1, 2,…,t
3: **PREDICTION:**
4: FOR i = 1, 2,…, N_s
5: Draw predicted particles from prior dynamics
6: Compute the colour histogram based on estimated position
7: END FOR
8: **MEASUREMENT:**
9: Calculate the Bhattacharyya distance, b_{dist}
10: Compute the weight of the particle based on Bhattacharyya distance, φ_c
11: Normalize the weight, $w_t^i = w_t^i(\sum_{i=1}^{N} w_t^i)^{-1}$
12: Calculate $\hat{N}_{eff}$
13: WHILE $\hat{N}_{eff} < N_s$
14: **RESAMPLING:**
15: IF $\hat{N}_{eff} < T_{thres}$
16: Choose the highest weight of the particle
17: Resample the discrete distribution based on highest weight particle
18: ELSE
19: Eliminate all low weight particles and remain the good particle
20: Regenerate the remaining of the particles
21: END ELSE
22: Compute the weight of the particle based on Bhattacharyya distance, φ_c
23: Check the effective sample size
24: END WHILE
25: **OUTPUT:**
26: Obtain the position of the target vehicle
27: **END FOR FRAME**

VI. RESULTS AND DISCUSSIONS

In this section, the results of vehicle tracking using the conventional particle filter resampling algorithm as shown in Fig. 2 will be compared to the results of vehicle tracking using an enhanced particle filter resampling algorithm as shown in Fig. 3. In both resampling algorithm, the particle size is initialized as 200 samples. Colour histogram is selected as the parameter to calculate the weight of the particles. A colour histogram with 8 × 8 × 8 bins RGB colour space will be computed. In this study, the colour feature has been chosen due to its ability to identify the

target vehicle identity during partial occlusion. In addition, when there are a small portion of the colour of the target vehicle is visible, the enhanced particle filter algorithm is able to locate the target vehicle. Since colour histogram in discrete form, the time required to process will be short.

Based on the results shown in Fig. 2 and Fig. 3, the crossing icon represents the location estimated by the particle filter algorithm. Meanwhile the solid box refers the bounding box of the target vehicle. The target vehicle position is determined by calculating the mean value of the coordinates estimated by each particle.

Referring to Fig. 2 and Fig. 3, it can be observed that the tracking results can be divided into four cases which are 'before occluded', 'partially occluded', 'fully occluded' and 'after occluded'. Comparing to the results shown in Fig. 2 and Fig. 3, it can be noticed that the results obtained by using the particle filter with adaptive resampling algorithm is much more promising.

In Case I, the target vehicle is not occluded by another static vehicle as shown in Frame 5 of Fig. 2 and Fig. 3. Based on the results, the conventional and enhanced particle filter resampling algorithms are able to track the target vehicle. This is because before the occlusion, the information of the target vehicle can be easily obtained without influences by another vehicle. Thus, the particle filter resamples as usual by eliminating the weak particles and replacing those unwanted particles with a new set of particles.

In Case II, the target vehicle is partially occluded by the static vehicle as shown in Frame 13 of Fig. 2 and Fig. 3. From the result shown, it is noticed that colour is an important feature that can be used to deal with partially occlusion incidents. In this case, the resampling algorithm is executed as the previous case because the information of the target vehicle still obtainable and not influenced by the static vehicle. Thus, the target vehicle is able to be tracked with both resampling algorithms.

In Case III, the target vehicle is almost fully occluded by the static vehicle as shown in Frame 16 of Fig. 2 and Fig. 3. Most of the information of the target vehicle is either lost or influenced by the information obtained for the static vehicle. Thus, the conventional resampling algorithm is failed to locate the target vehicle because the particles has been trapped at the static vehicle by getting the wrong information. Meanwhile, the improved resampling algorithm is able to track the target vehicle although only a small portion of the target vehicle colour is recognized.

In Case IV, the target vehicle occurs after the occlusion as shown in Frame 25 of Fig. 2 and Fig. 3. Based on the results obtained, the conventional resampling algorithm is unable to track the target vehicle. The conventional resampling algorithm failed in tracking the vehicle after occlusion is because the information of the target vehicle has been lost and replaced by the information of the static vehicle. Therefore, the target vehicle is considered lost track by using the conventional particle filter algorithm as

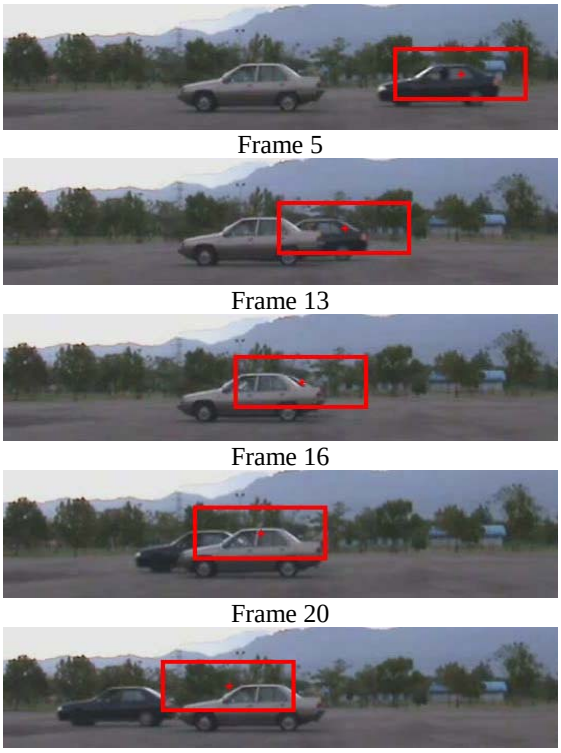


Figure 2. Results of vehicle tracking by using conventional resampling particle filter.

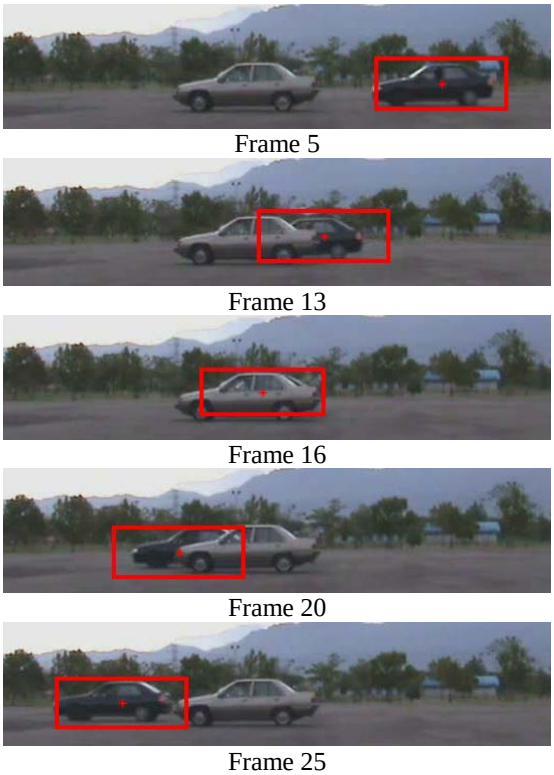


Figure 3. Results of vehicle tracking by using adaptive resampling particle filter.

shown in Frame 25 of Fig. 2. However, the enhanced particle filter with adaptive resampling is able to continuously track the target vehicle as shown in Frame 25 of Fig. 3. Although the information of the target vehicle will be influenced by the static vehicle after the overlapping, the improved resampling algorithm is capable of remaining the location of the highest weight particle. A new set of samples will be generated to gain back the information of the target vehicle. Hence, the target vehicle is able to be continuously tracked by the proposed algorithm effectively and efficiently.

Fig. 4 and Fig. 5 show the RMSE and the resampling process as well as the sample size versus the frame index. Based on Fig. 4 and Fig. 5, the frame index of 1 to 10 indicates the target vehicle is free from occlusion. The results shows that both of the resampling algorithms are able to track the target vehicle due to the low value of RMSE. However, in terms of computational time, the adaptive resampling is faster than the conventional resampling as shown in Fig. 5 due to the lesser resampling steps and number of resampling particles .

During the frame index of 11 to 14, the target vehicle is partially occluded by the static vehicle. From the result shown in Fig. 4, the RMSE for both resampling algorithms are almost the same. After frame index of 14, the number of resampled particles has been increased due to the influeces from the static vehicle as shown in Fig. 5. Frame index of 16 shows that the target vehicle is fully occluded by the static vehicle. From Fig. 4, the RMSE for the conventional algorithm is much more higher than the proposed algorithm. This is because the information of the target vehicle is influenced by the static vehicle and causes the conventional algorithm to an inaccurate result. Meanwhile, the tracking result for the enhanced resampling algorithm is still promissing with the low value of RMSE. In addition, Fig. 5 shows that the conventional algorithm will keep on resampling to locate the target vehicle and become an infinity loop. Thus, a maximum of 20 resampling steps has been set in order to terminate the infinity loop.

Frame index of 20 to 25 indicates the target vehicle is occurred after the occlusion. From the results shown in Fig. 4, the RMSE of the conventional algorithm is increasing due to the algorithm is still stuck at the static vehicle. The increasing of the RMSE means that the particle filter has diverged from tracking the target vehicle. However, RMSE for the enhanced algorithm is still remain at low level which means the visual tracker is still continuosly track the target vehicle. On the other hand, the number of resampling steps required for tha adaptive algorithm is still maintain at a low level. This means the adaptive algorithm can gain back the information of the target vehicle during and after the overlapping. However, the conventional algorithm is failed in tracking the target vehicle even though with the maximum of the resampling steps as shown in Fig. 5. The computational time taken by the

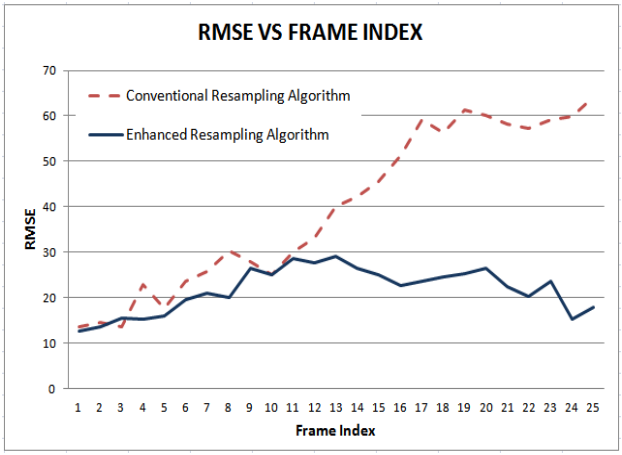


Figure 4. Graph of RMSE vs frame index.

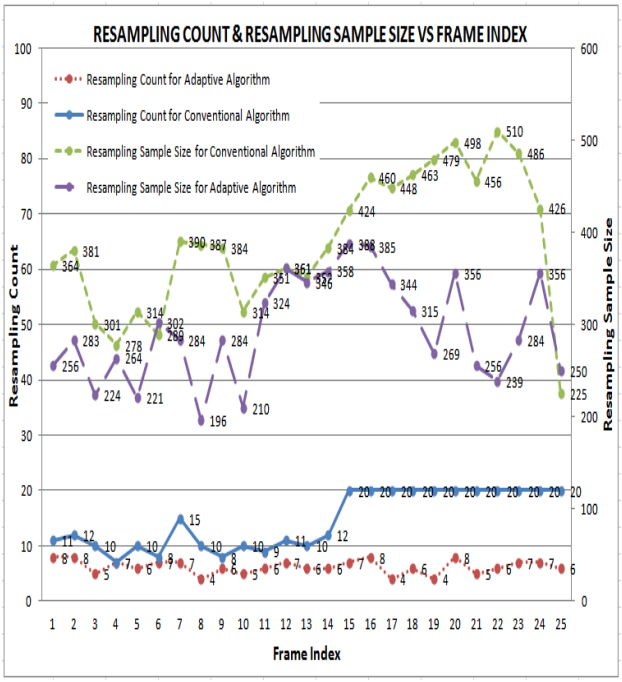


Figure 5. Graph of resampling count and resampling sample size vs frame index.

adaptive algorithm is shorter than the conventional algorithm due to the number of resampling steps and resampled particle sizes is lesser.

VII. CONCLUSION

As discussed in the previous sections, the accuracy of the particle filter algorithm will be diminished by the particle degeneracy. Conceptually, resampling can be used to avoid the particle degeneracy problem. Thus, an enhancement of the particle filter with adaptive resampling algorithm has been proposed for the purpose of tracking the overlapped vehicle. The implementation of the

adaptive algorithm in the particle filter approach is capable to ease the tracking difficulties of various occlusion incidents. The performance and robustness of the proposed algorithm is tested and assessed under various tracking conditions. It can be concluded that the adaptive resampling approach has improved the accuracy of the tracking results without compromising the computational time.

ACKNOWLEDGEMENT

The authors would like to acknowledge the financial assistance of the Ministry of Higher Education of Malaysia (MoHE) under Fundamental Research Grant Schemes (FRGS) No. FRG0220-TK-1/2010 and the University Postgraduate Research Scholarship Scheme (PGD) by Ministry of Science, Technology and Innovation of Malaysia (MOSTI).

REFERENCES

[1] F. Gustafsson, F. Gunnarsson, N. Bergman, U. Forsell, J. Jansson, R. Karlsson and P.J. Nordlund. "Particle Filters for Positioning, Navigation, and Tracking." *IEEE Transaction on Signal Processing*, vol. 50, no. 2. , pp. 425-437, 2002.

[2] W.Y. Kow, W.L. Khong, F. Wong, I. Saad and K.T.K. Teo. "Adaptive Tracking of Overlapping Vehicles via Markov Chain Monte Carlo with CUSUM Path Plot Algorithm." *In Proceedings of Computational Intelligence, Communication Systems and Networks*, 2011, pp. 253-258.

[3] P.L.M. Bouttefroy, A. Bouzerdoum, S.L. Phung and A. Beghdadi. "Vehicle Tracking using Projective Particle Filter." *In Proceedings of 6th IEEE International Conference on Advanced Video and Signal Based Surveillance*, 2009, pp. 7-12.

[4] J. Liu and P. Vadakkepat. "Interacting MCMC Particle Filter for Tracking Maneuvering Target." *Digital Signal Processing*, vol. 20, pp. 561-574, 2010.

[5] J. Li, X. Lu, L. Ding and H. Lu. "Moving Target Tracking via Particle Filter Based on Colour and Contour Features." *In Proceedings of 2nd International Conference on Information Engineering and Computer Science*, 2010, pp.1-4.

[6] J. Czyz, B. Ristic and B. Macq. "A Color-based Particle Filter for Joint Detection and Tracking of Multiple Objects." *In Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing.* 2005, pp. 217-220.

[7] C.J. Yang, R. Duraiswami and L. Davis. "Fast Multiple Object Tracking via a Hierarchical Particle Filter." *In Proceedings of 10th IEEE International Conference on Computer Vision*, 2005, pp. 212-219.

[8] M.S. Arulampalam, S. Maskell, N. Gordon and T. Clapp. "A Tutorial on Particle Filter for Online Nonlinear/Non-Gaussian Bayesian Tracking." *IEEE Transaction on Signal Processing*, vol. 50, no.2, pp. 174-188, 2002.

[9] H.P. Liu, F.C. Sun, L.P. Yu and K.Z. He. "Vehicle Tracking using Stochastic Fusion-based Particle Filter." *In Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2007, pp. 2735-2740.

[10] K. Nummiaro, E. Koller-meier and L.V. Gool. "Colour Features for Tracking Non-Rigid Objects." *Special Issue on Video Surveillance Chinese Jornal of Automation*, vol. 29, pp. 345-355, 2003.

[11] T. Wada, F. Huang and S. Lin. "Visual Tracking Using Particle Filters with Gaussian Progress Regression." *Springer-Verlah Berlin Heidelberg 2009. PSIVT 2009, LNCS 5414*, 2009, pp. 261-270.

[12] X.Y. Fu and Y.M. Jia. "An Improvement on Resampling Algorithm of Particle Filter." *IEEE Transaction on Signal Processing*, vol. 58, no.10, pp. 5414-5420, 2010.

[13] A.D. Bagdanov, A.D. Bimbo, F. Dini and W. Numziati. "Adaptive Uncertainty Estimation for Particle Filter-based Trackers." *In Proceedings of 14th International Conference on Image Analysis and Processing*, 2007, pp. 331-336.

[14] T. Zhang, S.M. Fei, X.D. Li and H. Lu. "An Improved Particle Filter for Tracking Color Object." *In Proceedings of International Conference on Intelligent Computation Technology and Automation*, 2008, pp. 109-113.

[15] M.S. Khalid, M.U. Ilyas, M.S. Sarfaraz and M.A. Ajaz. "Bhattacharyya Coefficient in Correlation of Gray-Scale Objects." *Journal of Multimedia*, vol.1 no. 1, pp. 56-61, 2006.

Effect of Partially Shaded Conditions on Photovoltaic Array's Maximum Power Point Tracking

Chia Seet Chin, Prabhakaran Neelakantan, Soo Siang Yang,
Bih Lii Chua, Kenneth Tze Kin Teo
Modelling, Simulation and Computing Laboratory
School of Engineering and Information Technology
Universiti Malaysia Sabah, Kota Kinabalu, Malaysia.
msclab@ums.edu.my, ktkteo@ieee.org

***Abstract* - Maximum power point tracking algorithm is widely implemented in photovoltaic system to maximize the PV array output power. In general, Perturb and Observe (P&O) is simple thus being selected to continuously track the array maximum power point (MPP). Under uniform solar irradiance, PV array characteristic is non-linear and consisting only one MPP along the functional operating voltage. However, when the PV array is partially shaded, the P-V characteristic becomes more complex with multiple MPPs. The occurrence of multiple MPP might cause the PV array to be trapped at the local MPP. At this operating condition of local MPP, PV array will generate lesser output power. In this study, the performance of PV array is explored especially when each PV module is at 30% and 70% shaded conditions. Simulation results show that PV array at absolute MPP can generate greater output power with the largest increased by 74.6% hence achieving higher power efficiency.**

Keywords - PV array, partially shaded conditions, maximum power point, perturb and observe

I. INTRDUCTION

The world primary energy demand is expected to increase 1.7% per year from the year 2002 and expending more than 50% at 2030 [1]. It is projected the main global energy consumption will be continuously dominated by fossil fuel. Due to the awareness of global warming and climate change, nations are concerned of the planet's carbon emissions and fossil fuel used. Therefore, renewable energy resources such as solar energy will play a significant role in the world energy in the upcoming future [2].

Solar energy is the sun's radiant energy where the radiate amount is enormous. Solar energy can be converted directly into electrical energy via solar cell and the phenomenon is commonly described as photovoltaic (PV) effect [3]. Solar energy has gained the world's interest as it is a promising option of renewable energy. The conversion of solar energy to electrical energy is static, quite, free of moving object and does not have any negative impact to the environment. Therefore, the annual growth of the world PV industry has reached an average of 30% during the past decade [4].

Commercial PV module is configured from series, parallel or combination of series and parallel connection of PV cells. Solar irradiance, cell temperature, tilt angle shaded condition and the operating condition are among the factors that will influence the output current and power characteristics of the PV module. However, the electricity generation is mainly affected by the level of solar irradiance. The amount of incident light to the PV module will determine the total generation charge carrier hence the generated current in the PV module [5, 6]. In general, PV array is formed by a couple of PV modules to obtain larger output power.

II. LITERATURE REVIEW

In [2], [3] and [4], the authors presented the non-linearity of the *I-V* and *P-V* characteristics when the PV system is under uniform illumination of sunlight. Generally, PV system has a unique operating point named as maximum power point (MPP) which will produce maximum power at the optimum voltage and current. The authors in [4] proved that the MPP is varied according to the amount of solar irradiance as the PV characteristics diverse as the changing of the atmospheric conditions. Cell temperature is one of the factors discussed in [3] which will influence the position of MPP. However, [3] and [4] did not discuss the effect of partially shaded condition on the PV system.

In [3] and [7], the authors proposed maximum power point tracking (MPPT) algorithm to identify the MPP in the uniform illuminated PV system. According to [7], Perturb and Observe (P&O) method is commonly used for its simplicity and low cost implementation. P&O is able to maximize the power generation of the PV system by continuously tracking the MPP regardless of the changing of the atmospheric conditions. However, the research is only limited to the uniform illuminated conditions.

The studies in [8], [9] and [10] show that when the PV array is under partially shaded conditions (PSC), the array characteristics become more complex with multiple MPPs. PSC is defined as the circumstance where one or more of the PV modules in the array received less amount of solar irradiance. The existence of multiple MPPs reduces the effectiveness of MPPT algorithm since the power losses of

PV system under partially shaded conditions can reach as high as 70% [8].

In this paper, the effect of the partially shaded conditions on PV array's MPPT will be studied. The occurrence of multiple MPPs generally might cause the PV array to be trapped at the local MPP. At this operating condition, PV array will generate lesser output power. On the other hand, if the PV array is operated at the absolute MPP, the array can generate greater output power. Therefore, PV array which is operated at the absolute MPP can achieve higher power efficiency compared to the trapped local MPP.

III. MODELLING OF PV SYSTEM AND COMPUTATIONAL OF MPPT

This section described the mathematical model of the basic element solar cell. The mathematical model of solar cell can be further implemented for the PV module and PV array modelling. In addition, the computational of Perturb and Observe (P&O) will be discussed.

A. Mathematical Model of PV Module

Solar cell is the basic element in PV module and it can be represented by one-diode model as shown in Fig. 1. The equivalent circuit of ideal PV cell consists of a photo current source and a diode [11]. In order to model the PV characteristics more practically, additional parameters such as equivalent parallel resistor, R_p and equivalent series resistor, R_s should be considered. R_p existed in the solar cell mainly due to the p-n junction leakage current whereas the occurrence of R_s is caused by the contact resistance within the semiconductor layer of the metal base [12].

The *I-V* characteristic of diode D_m can be described by the Schockley diode equation as shown in (1) where I_{Dm} is the current that flow through diode D_m, I_0 is the diode D_m reverse bias saturation current, V_{Dm} is the voltage across the diode D_m, n is the ideality factor of the diode D_m and V_T is the thermal voltage.

$$I_{D_m} = I_0\left[\exp\left(\frac{V_{D_m}}{nV_T}\right)-1\right] \tag{1}$$

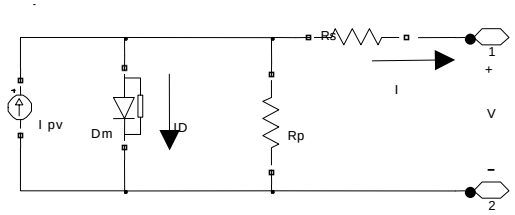


Figure 1. One-diode model.

The thermal voltage, V_T is shown in (2) where k is the Boltzman constant ($1.3806503 \times 10^{-23}$ J/K), T is cell's operating temperature in degree Kelvin and q is the electron charge ($1.60217646 \times 10^{-19}$ C).

$$V_T = \frac{kT}{q} \tag{2}$$

Based on the general model of the solar cell as in Fig. 1, the *I-V* characteristic of the solar cell can be further described as in (3), where I is the terminal current of the solar cell, I_{pv} is the light-generated current of the solar cell, V is the solar cell terminal voltage, R_s is the equivalent series resistance and R_p is the equivalent parallel resistance.

$$I = I_{pv} - I_0\left[\exp\left(\frac{V+IR_S}{nV_T}\right)-1\right]-\left(\frac{V+IR_S}{R_p}\right) \tag{3}$$

In general, PV module consists of several identical solar cell connected in series to provide higher operating voltage for the PV module [11]. PV module which consists of n series connected solar cell can be modelled with modification in (2). The thermal voltage in (2) will be multiplied by the number of series solar cell, n.

B. PV Array with By-Pass Diode

PV array is formed to have larger output power generation. PV array which is formed by series connected PV module has larger functional operating voltage [13]. Fig. 2 shows the forming of PV array by three identical PV modules connected in series.

The modelling of the characteristic of PV array is based on superposition of each individual PV module characteristic. Theoretically, PV array will produce same amount of generating current as the PV module which is operated under the same environmental conditions on the other hand, the PV array provides greater output voltage. Referring to Fig. 2, the array will gain three times operating voltage compared to the PV module.

By-pass diode is externally coupled to each PV module to prevent hot spot formation in the PV array. During

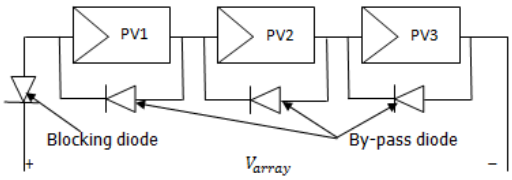


Figure 2. PV array consists of three identical PV modules.

uniform solar illumination, the entire PV modules will generate same amount of current. However, the shaded module will generate less current during PSC. The shaded module will cause the overall current generation by the entire PV array being limited. The solar cells will be reverse biased due to the mismatched effect and dissipate power in form of heat. When the hot spots exceed the maximum power which can be sustained by the PV cells, it will cause permanent damage to the PV module and the array will be open circuited causing power interruption to the users [14]. With by-pass diode, the excess current by the un-shaded PV module is allowed to flow through the external diode, preventing the PV module from being limited by shaded module and avoiding the destructive effect caused by hot spot formation in the PV array.

Blocking diode is externally connected to the string of series connected PV modules. The function of blocking diode is to prevent the reverse flow of current particularly from the secondary power sources such as the lead-acid battery during the absence of sun light.

C. *Perturb and Observe*

Perturb and Observe (P&O) is implemented to control the PV array and optimize the power generation under varying atmospheric conditions. P&O has simple control structure where it needs a few measured parameters to perform maximum power tracking [15]. P&O operates periodically for perturbing the control variable and comparing the instantaneous PV output power at each sampling interval.

P&O is initiated by applying perturbed voltage to change the operating voltage of the PV array. Subsequently, module parameters V and I at the present state, k and previous state, $k-1$ are measured. The module output power, P can be calculated by multiplying the parameters V and I. Based on the comparison of present and previous state of the operating voltage and output power, P&O will make decision either to shift the PV array for larger or lower operating voltage. Eventually, the PV array will be operated at MPP where PV array generates maximum power. Fig. 3 shows the operation of P&O.

Fig. 4 shows the P-V characteristic of the PV module under 600W/m^2 solar irradiance with the MPP tracking process of P&O. The shifting of the instantaneous operating voltage to approach the MPP in Fig. 4 is based on the P&O operation as in the flowchart in Fig. 3. There are total of four possible circumstances which will affect the direction of the MPP tracking.

Case I where $P_k > P_{k-1}$ and $V_k > V_{k-1}$ is illustrated as tracking path α in Fig. 4. Increasing of PV module operating voltage will lead to larger PV power. Hence, P&O will apply an increment of perturbed voltage, ΔV to the present PV module operating voltage until the MPP is successfully determined.

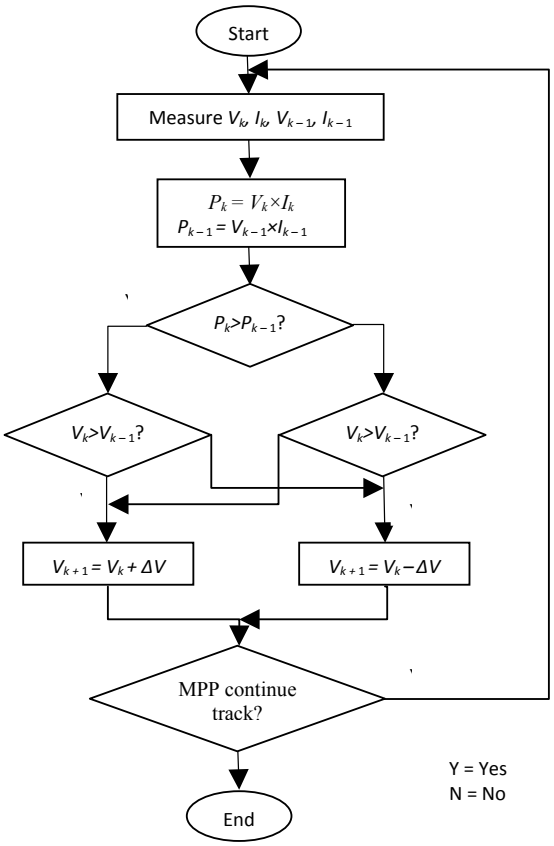


Figure 3. Flowchart of P&O operation.

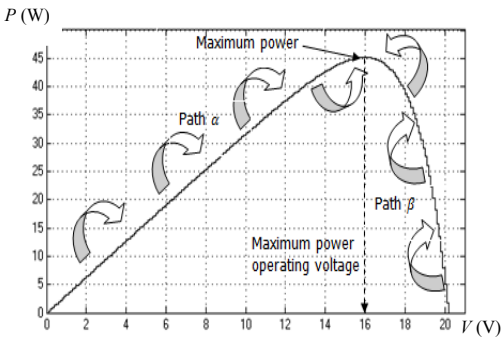


Figure 4. P&O tracking path.

Case II where $P_k > P_{k-1}$ and $V_k < V_{k-1}$ is described as tracking path β in Fig. 4. At this condition, the PV module output power will increase as the decreasing of PV operating voltage. Thus, P&O will take action to reduce the perturbed voltage, ΔV to the present PV module operating voltage. While the condition is unchanged, the decrement of ΔV will be continued until the MPP is identified.

Case III involves $P_k < P_{k-1}$ and $V_k > V_{k-1}$ which is represented by the tracking path β in Fig. 4. In this case, the PV power decreases as the PV operating voltage increases.

Thus, P&O applies perturbed voltage, ΔV to shift the PV module's operating voltage to lower value.

Case IV is illustrated as $P_k<P_{k-1}$ and $V_k<V_{k-1}$ and it is described by tracking path α in Fig. 4. It can be observed that when the PV operating voltage is reduced, the output power will be decreased. Thus, P&O will apply perturbed voltage, ΔV to the PV module and operate at larger voltage.

P&O will continue to track MPP even the optimal voltage point has been successfully achieved. The continuous MPP tracking is vital to discover the next MPP for maximum power gaining especially when the PV system is under rapid changes of environmental conditions.

IV. SIMULATION OF PV SYSTEM

Modelling and simulation of PV system has been carried out in MATLAB-SIMULINK. The commercial PV module, SHARP NE-80E2EA multi-crystalline silicon with 80W is selected as the reference model for the modelling. SHARP NE-80E2EA consists of 36 series connected solar cell with 21.3V open circuited voltage and 5.16A short circuit current.

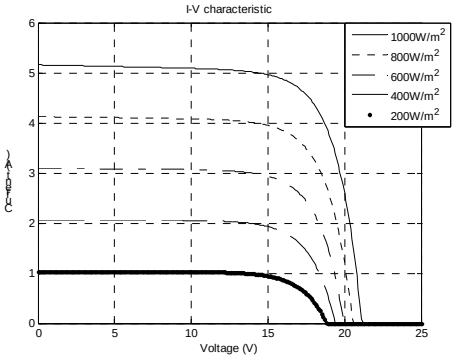
A. Uniform Illuminated Conditions

Fig. 5 shows the *I-V* and *P-V* characteristics of PV modules at various solar irradiances. The characteristics describe the output current and power of the PV module within the functional operating voltage. The amount of the illuminated solar irradiance affects the output current generation. As the solar irradiance increased, PV cells in the module are able to release more electrons thus generating larger current. Hence, the output power is increased as the growing of solar irradiance. The *P-V* characteristic illustrates non-linearity behaviour with appearance of one MPP. At 1000W/m^2 solar irradiance, the MPP is located at 17.1V with the power generation of 80W while at 200W/m^2 solar irradiance, the MPP is relocated to 15.1V with the power generation of 16W.

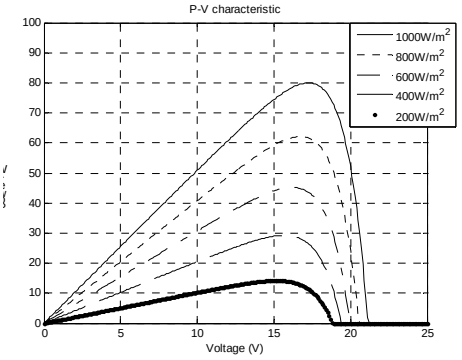
Fig. 6 shows the *I-V* and *P-V* characteristics of the PV array. In this study, PV array is formed by three identical PV module connected in series. Therefore, PV array has three times larger operating voltage compared to the single PV module where in Fig. 6, the functional operating voltage range of PV array has three times of the voltage magnitude of PV module in Fig. 5. However, the series connected PV modules will not amplify the current generation. The current generated by thePV array is the same as PV module. Nevertheless, the array power generation will be three times greater than the output power produced by single PV module.

B. Partially Shaded Conditions

PV array under uniform illuminated conditions has non-linear characteristic with the occurrence of one MPP in the

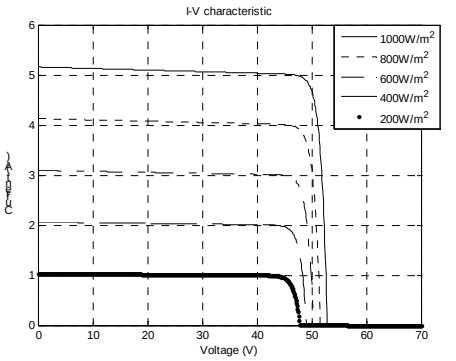


(a)

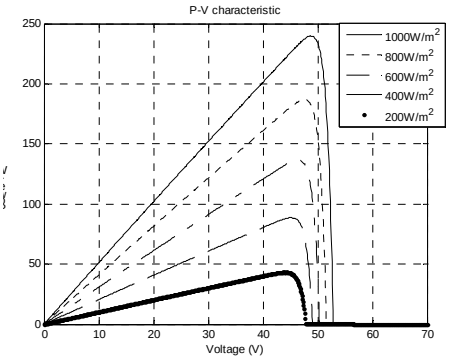


(b)

Figure 5. The (a) *I-V* and (b) *P-V* characteristics of PV module.



(a)



(b)

Figure 6. The (a) *I-V* and (b) *P-V* characteristics of PV array.

P-V curve. However, when the PV array is under partially shaded conditions, the *P-V* characteristic becomes more complex. Multiple MPPs occur in the *P-V* characteristic due to the mismatched current generation by the PV array. Fig. 7 and Fig. 8 illustrate the characteristics of PV array when one of the PV modules in the PV array is under shaded condition of 20% and 50% respectively. There are two MPPs in the *P-V* characteristic as shown in Fig. 7. It can be observed that PV array which is exposed to 20% partial shading has a local MPP of 34V and an absolute MPP of 48V. At local MPP, PV array will have approximately 170W power generation. However, if the PV array is operated at the absolute MPP, the power generation can be boost up to approximately 190W, 11.8% more than the local MPP.

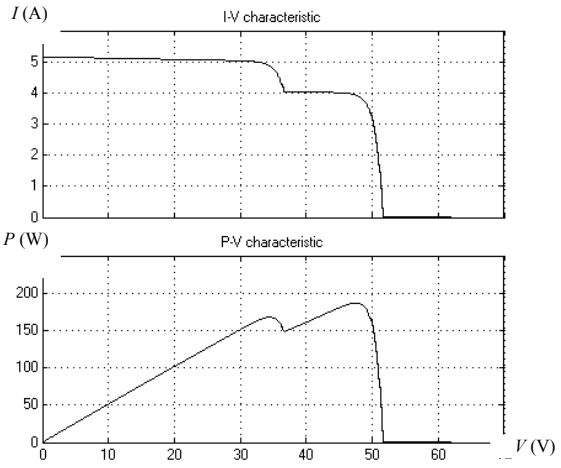


Figure 7. Characteristics of PV array where one of PV module shaded 20%.

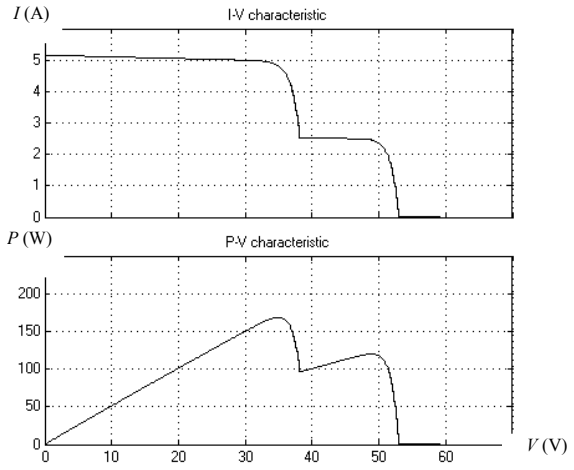


Figure 8. Characteristics of PV array where one of PV module shaded 50%.

Considering Fig. 8, where PV array is exposed to 50% partially shaded condition, the array presents a local MPP at operating voltage of 50V and an absolute MPP at 34V. PV array can generate output power of 120W in the local MPP operating condition. However, the efficiency of PV array will be enhanced if the PV array shifted the operation to the absolute MPP, where PV array can generate 41.7% larger power than the local MPP.

Fig. 9 shows the characteristics of PV array when two of the PV modules are shaded 50%. It can be observed that there are two MPPs in the *P-V* curve which are located at the operating voltage of 17V and 50V respectively. In this situation, PV array which is operated at 17V is considered trapped at the local MPP. PV array at the trapped local MPP is only able to generate output power of approximately 80W. On the other hand, if the PV array is able to shift the operating voltage beyond the trapped local MPP and operated at 50V, the output power will be increased to 130W, 160% larger than the local MPP.

Table I summarizes the location of both local and absolute MPP when the PV array is under various partially shaded conditions. The effect of PV array power generation at absolute MPP compared to local MPP is also included in the table.

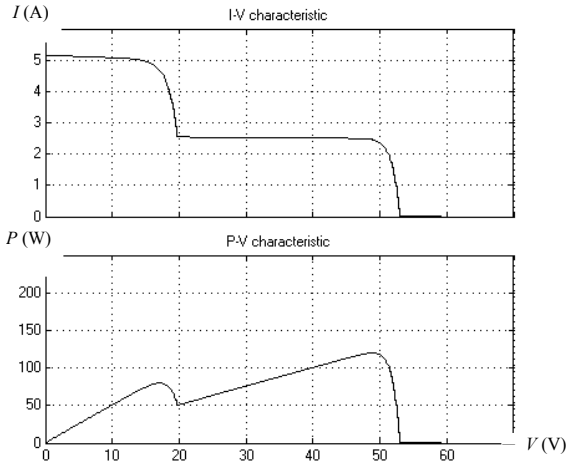


Figure 9. Characteristics of PV array where two of PV modules shaded 50%.

Table I. MPPs and the effect to the PV array power generation.

Number of PV modules shaded (% shaded)	MPPs operating voltage (V)		Effect of power generation (Absolute MPP operating condition)
	Local	Absolute	
One (20%)	34	48	Increased of 11.8%
One (50%)	34	50	Increased of 41.7%
Two (50%)	17	50	Increased of 160%

When the PV array is exposed to various partial shading, multiple local MPPs might be occurred in the *P-V* characteristic. Fig. 10 shows the PV characteristics when each PV modules in PV array is under 30% and 40% shaded conditions, whereas Fig. 11 shows the PV characteristics when each PV modules in the array is under 30% and 70% shaded conditions.

It can be noticed from Fig. 10, there are two local MPPs and one absolute MPP. The absolute MPP is located at the operating voltage of 47V. At local MPP of 17V, PV array generates output power of approximately 80W. There is another local MPP beyond the local MPP of 17V, which is situated at the operating voltage of 32V. The PV array generates greater output power, 110W if it is operated at local MPP of 32V. However, at absolute MPP operating condition, the PV array produce the highest output power generation, which is 140W.

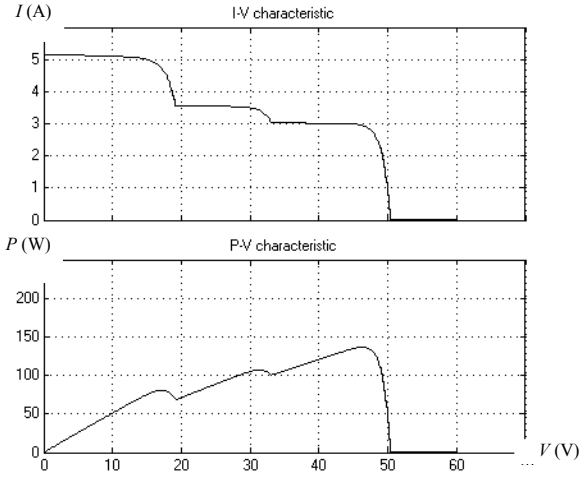


Figure 10. Characteristics of PV array where PV modules shaded 30% and 40%.

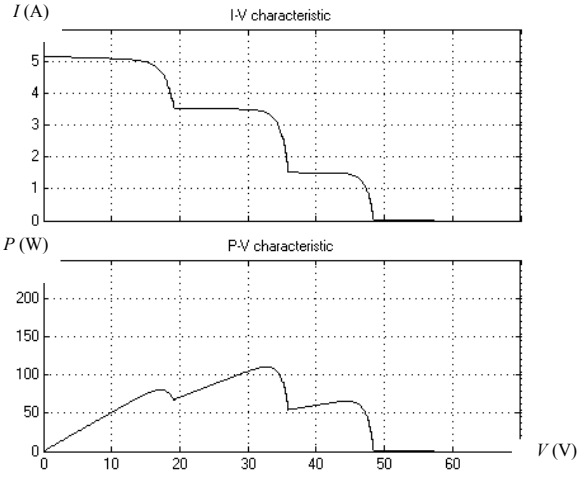


Figure 11. Characteristics of PV array where PV modules shaded 30% and 70%.

There are three MPPs being detected in Fig. 11. The MPPs are situated at the operating voltage of 17V, 32V and 47V respectively. Unlike Fig. 10, Fig. 11 shows that the absolute MPP located at the operating voltage of 32V. At this operating condition, the PV array is able to generate maximum power of 110W. PV array will have smaller power generation if it is operated at the local MPPs of 17V and 32V. Thus, the efficiency of PV array operated at the local MPPs will be reduced.

V. RESULTS

Fig. 12, Fig. 13 and Fig. 14 shows the tracking operation of MPPT on partially shaded PV array in order to maximize the output power generation. The PV array is under partially shaded condition where one of module is shaded 30% while another PV module is shaded 70%. Those corresponding characteristics can be referred in Fig. 11.

VI. DISCUSSION

P&O is one of the most common MPPT methods used to track the MPP of PV array for maximum power gaining purpose. In general, P&O performs well when the PV module or PV array is under uniform illuminated solar irradiance. However, partially shaded conditions can significantly affect the operation of the P&O.

Referring to Fig. 12, PV array is initially operated at 10V. At this condition, PV array is able to generate output power of 50W. P&O is implemented in the PV system to regulate the operating voltage of the PV array for greater output power. By comparing the operating voltage and the measured power at two sampling instants, P&O decides to

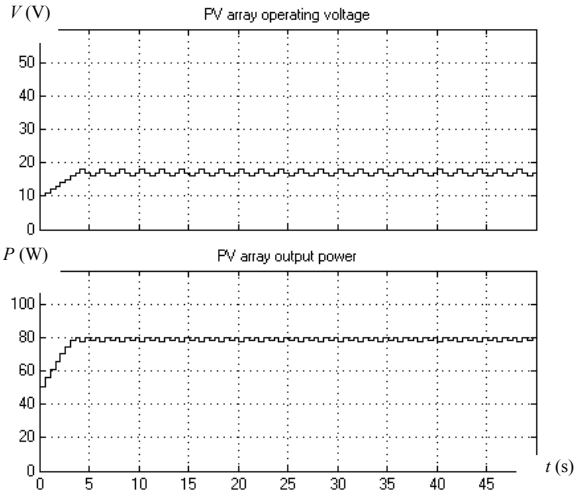


Figure 12. PV array trapped at local MPP of 17V.

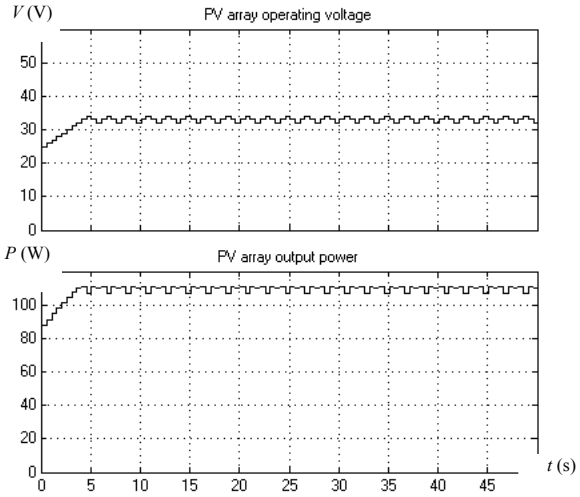


Figure 13. PV array operated at absolute MPP of 34V.

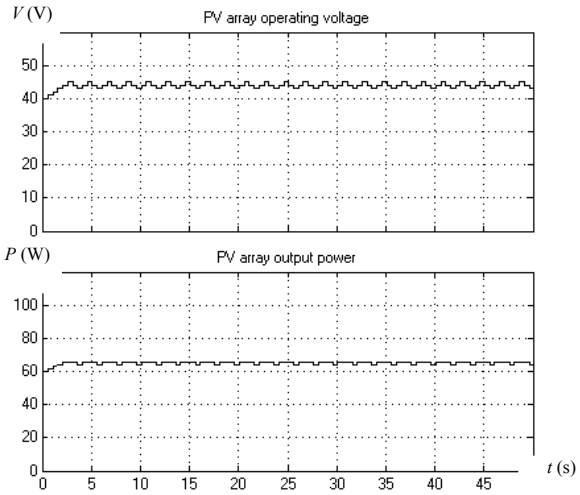


Figure 14. PV array trapped at local MPP of 47V.

shift the array for larger operating voltage. Thus, the PV array operating voltage is increased until it reached MPP at 17V. At this operating condition, the PV array is able to generate output power of 80W. The output power has been increased by 60% compared to the initial operating condition.

The operation of P&O is based on comparison of parameters at two sampling interval. Thus, P&O instructs the array to maintain the operating voltage at around 17V. Result in Fig. 12 shows that voltage fluctuation has occurred. This is because P&O attempts to track another MPP during the next perturbation cycle.

The MPPT might be able to initiate the tracking process from another initial voltage point. Referring to Fig. 13, the initial operating voltage of PV array is 25V. At this situation, PV array has the power generation of approximately 82W. To obtain maximum power, P&O increase the operating voltage of PV array, hence achieving higher power generation. The voltage increment process is continued until it reaches MPP at 32V. PV array is able to generate output power of approximately 110W. Compared to the initial voltage, the power generation has been increased by 34.1%.

In Fig. 14, P&O start the tracking process from 40V. At 40V, PV array is able to generate output power of 60W. However, P&O performs tracking process and shifts the PV array to a larger operating voltage. At the MPP of 42V, the PV array is able to generate higher output power of approximately 63W, an increase of 5% compared to the initial operating voltage.

During partially shaded conditions, PV array presents a more complex *P-V* characteristic with multiple MPPs. The occurrence of multiple MPP can cause the PV array to be trapped at the local MPP. Fig. 12 and Fig. 14 are the simulation results which show that the PV array has been trapped at the local MPP. Although P&O has successfully track the MPP, the output power generation at the trapped local MPP of 17V and 47V is less than the absolute MPP of 32V. At absolute MPP operating condition, PV array can generate output power of 110W, which is 37.5% and 74.6% greater that the power generation of the PV array at local MPP of 17V and 47V respectively. Hence, PV array which is operated at absolute MPP can achieve higher efficiency of PV system than the trapped local MPP. Table II summarises the power improvement of PV array operated at absolute MPP compared to the trapped local MPPs.

VII. CONCLUSION

The power efficiency of PV system can be improved by implementing MPPT algorithm to the system to track the optimal MPP. P&O is selected to perform the tracking because of its ease and low cost implementation. PV array at uniform solar irradiance presents non-linear characteristic. However, the characteristic is simple with only one MPP in the *P-V* curvature. Partially shaded conditions cause the *P-V* characteristic become more complex with multiple MPPs. This situation might lead the PV array to be trapped at the local MPP and reduce the power generation. Results show that PV array which is

Table II. Effect of different shaded conditions to the PV array power generation.

PV array where each module under 30% and 70% shaded conditions		
MPPs operating voltage (V)		**Effect of power generation (Absolute MPP operating condition)**
Local	**Absolute**	
17	32	Increased of 37.5%
47	32	Increased of 74.6%

operated at the absolute MPP can achieve the largest power generation as high as 74.6% improvement than the trapped local MPP. Therefore, it can be concluded that PV array at the absolute MPP operating condition will achieve higher efficiency compared to the trapped local MPP.

ACKNOWLEDMENT

The authors would like to acknowledge the financial assistance of University Postgraduate Research Scholarship Scheme (PGD) by Ministry of Science, Technology and Innovation Malaysia (MOSTI).

REFERENCES

[1] M.R. Islam, R. Saidur, N.A. Rahim, and K,H, Solangi, "Usage of solar energy and its status in Malaysia," Engineering e-Transaction, vol. 5, no. 1, pp. 6-10, 2010.

[2] H.L. Tsai, "Insolation-oriented model of photovoltaic module using MATLAB/SIMULINK," Solar Energy, vol.84, issue 7, pp. 1318-1326, 2010.

[3] C.S. Chin, P. Neelakantan, H.P. Yoong, and K.T.K. Teo, "Fuzzy logic based MPPT for photovotaic modules influenced by solar irradiance and cell temperature," Proceedings of 13th International Conference on Computer Modelling and Simulation, Cambridge, United Kingdom, 2011, pp. 376-381.

[4] Syafaruddin, E. Karatepe, and T. Hiyama, "Polar coordinate fuzzy controller based real-time maximum-power point control of photovoltaic system," Renewable Energy, vol. 34, issue 12, pp. 2597-2606, 2009.

[5] H. Patel, and V. Agarwal, "MATLAB-based modeling to study the effects of partial shading on PV array characteristics," IEEE Transaction on Energy Conversion, vol. 23, no. 1, pp. 302-310, 2008.

[6] M.G. Villalva, J.R. Gazoli, and E.R. Filho, "Modeling and circuit-based simulation of photovoltaic arrays," Proceedings of Power Electronics Conference, Bonito, Brazil, 2009, pp. 1244-1254.

[7] N. Femia, G. Petrone, G. Spagnuolo, and M. Vitelli, "Optimization of perturb and observe maximum power point tracking method," IEEE Transaction on Power Electronics, vol. 20, issue 4, pp. 963-973, 2005.

[8] Y.H. Ji, D.Y. Jung, J.G. Kim, J.H. Kim, T.W. Lee, and C.Y. Won, "A real maximum power point tracking method for mismatching compensation in PV array under partially shaded conditions," IEEE Transaction on Power Electronics, vol. 26, issue 4, pp. 1001-1009, 2011.

[9] H. Patel, and V. Agarwal, "Maximum power point tracking scheme for PV systems operating under partially shaded conditions," IEEE Transaction on Industrial Electronics, vol. 55, no. 4, pp. 1689-1698, 2008.

[10] S.R. Chowdhury, and S. Hiranmay, "Maximum power point tracking of partially shaded solar photovoltaic arrays," Solar Energy Materials and Solar Cells, vol. 94, issue 9, pp. 1441-1447, 2010.

[11] M.C.D. Piazza, and G. Vitale, "Photovoltaic field emulation including dynamic and partial shadow conditions," Applied Energy, vol. 87, issue 3, pp. 814-823, 2010.

[12] M.G. Villalva, J.R. Gazoli, and E.R. Filho, "Comprehensive approach to modeling and simulation of photovoltaic arrays," IEEE Transactions on Power Electronics, vol. 24, no. 5, pp. 1198-1208, 2009.

[13] Y.J. Wang, and P.C. Hsu, "Analytical modelling of partial shading and different orientation of photovoltaic modules," IET Renewable Power Generation, vol. 4, issue 3, pp. 272-282, 2010.

[14] S. Silvestre, A. Boronat, and A. Chouder, "Study of bypass diodes configuration on PV modules," Applied Energy, vol. 86, issue 9, pp. 1632-1640, 2009.

[15] N.A. Ahmed, and M. Miyatake, "A novel maximum power point tracking for photovolatic application under partially shaded insolation conditions," Electric Power Systems Research, vol.78, issue 5, pp. 777-784, 2008.