

Object Tracking for Laparoscopic Surgery Using the Adaptive Mean-Shift Kalman Algorithm

Vera Sa-Ing, Saowapak S. Thongvigitmanee, Chumpon Wilasrusmee, and Jackrit Suthakorn

Abstract—In this paper, we propose the adaptive mean-shift Kalman tracking algorithm based on the mean-shift algorithm and the Kalman filter for tracking a laparoscopic instrument in laparoscopic surgery. With an iterative update of the target candidate in the mean-shift process, the proposed algorithm has improved the tracking performance over a typical mean-shift algorithm. In addition, the Kalman filter is employed to enhance the chance of tracking accuracy, especially when the object disappears from the scene. We tested the tracking performance of our proposed algorithm through simulated videos and real laparoscopic surgery videos. From all experimental results, the proposed algorithm can locate the target object correctly even when the size and the shape of the target have been changed. In the difficult situation when the target is hiding behind an obstacle, this algorithm can still track the target object correctly after it comes out. In addition, the proposed algorithm can be applied to locate different types of laparoscopic instruments.

Index Terms—Mean-Shift algorithm, Kalman filter, object tracking, laparoscopic surgery

I. INTRODUCTION

The development of surgery is moving toward minimization or elimination of incision, which is known as minimally invasive surgery (MIS) [1]. Patients can receive many benefits from MIS. For example, blood loss can be reduced, small incision results in reducing pain and shortens recovery time; and exposure of internal organs to possible external contaminants decreases the risk of infections. Laparoscopic surgery, a type of MIS, is performed with several laparoscopic instruments and a laparoscope which is a telescopic rod lens system connecting to a CCD camera. A fiber optic cable system is used to connect a light source to illuminate an operative field, which is viewed through a laparoscope to display a 2D image. The position of the laparoscope is always changed according to the instruction of the primary surgeon. In addition, the surgeon requires a lot of skills to operate the laparoscopic instrument.

Many researchers have proposed different methods to track the laparoscopic instrument. Omote *et al.* [2] presented the color tracking algorithm to control a robotic laparoscope instead of using human; however, this method cannot track many types of instruments. Casals *et al.* [3] introduced

feature tracking algorithm based on shape information of a surgical instrument; however, it works only with a specific surgical instrument. Lee *et al.* [4] proposed a color and shape tracking algorithm by using the contour of the surgical instrument. This algorithm worked well in the normal situation, but not when the instrument is blocked by some obstacles. Wei *et al.* [5] presented a simple algorithm for tracking target features. However, this algorithm is based on the artificial color marks attached to a surgical instrument, but there are many disadvantages, such as sterilization of the mark on the surgical instrument and its convenience in the real practice.

Due to some limitations of previous methods, in this paper, we propose a new object tracking algorithm to track the surgical instrument called the adaptive mean-shift Kalman algorithm [6], which is based on the mean-shift algorithm [7] and the Kalman filter [8]. In this technique, the size of the target candidate can be adjusted during tracking processes to increase the chance of tracking. Different scenarios of simulated videos and real laparoscopic videos were tested with the proposed algorithm, in this paper. Furthermore, the proposed algorithm is intended to use for controlling our new laparoscopic-holder assistant robot [9-10] and tracking the tip's instrument in laparoscopic surgery.

II. OBJECT TRACKING SYSTEM

In general, object tracking can be divided into two parts: *Target Representation and Localization*, and *Filtering and Data Association* [11]. The *Target Representation and Localization* part is the main process for locating and tracking a target object. It consists of three different techniques: point tracking, silhouette tracking, and kernel tracking. In this study, we will focus on kernel tracking, specifically the Mean-Shift algorithm, to locate the target object. The mean-shift algorithm can locate both rigid and non-rigid objects, as well as varying-size objects. *Filtering and Data Association* is a supplementary process utilizing the result of the target representation and localization process to improve object tracking capability to overcome some difficult tasks, such as a situation when the object is blocked by some obstacles. Examples of filtering algorithms are Kalman Filters and Particle Filters. In this part, the Kalman filter is the suitable estimation state of one target object, which is our main focus.

A. Mean-Shift Algorithm

A Mean-Shift algorithm [7] is an iterative process to locate the target object by maximizing the similarity function. The similarity function will be compared between the target model, \hat{q} , and the target candidate, $\hat{p}(y)$. The target model

Manuscript received October 12, 2011, revised October 13, 2011.

Vera Sa-Ing and Jackrit Suthakorn are with the Department of Biomedical Engineering, Mahidol University, Thailand (e-mail: jack_rotor@hotmail.com).

Saowapak S. Thongvigitmanee is with Image Technology Lab, National Electronics and Computer Technology Center, Thailand

Chumpon Wilasrusmee is with the Department of Surgery, Ramathibodi Hospital, Thailand

and the target candidate are represented by a small elliptical or rectangular area in the frame. Then pixel values in the region of interest (ROI) are used for calculating the target model and target candidate histograms. The main algorithm consists of five steps when computing each frame of video sequences. The result of processing in the current frame is the target object which is the target model in the mean-shift algorithm. The five steps are performed as shown in Fig. 1.

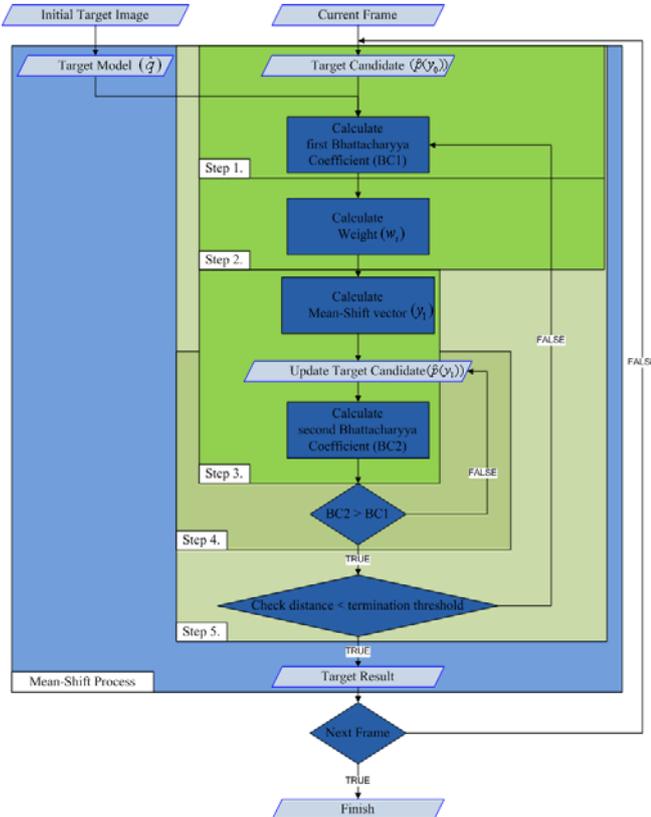


Fig. 1. The overall process of Mean-Shift algorithm

Step 1: In the first frame, we need to initialize the target model, $\{\hat{q}_u\}_{u=1\dots m}$, to be its distribution and y_0 to be the center location of the target model in the current frame. The target model can be computed by the following equation:

$$\hat{q}_u = C \sum_{b(x_i)=u} k(\|x_i\|^2) \quad (1)$$

where x_i is the normalized pixel value at the i th pixel of the target model area, $b(x_i)$ is a color value at pixel x_i which depends on the m -bin histogram, C is a normalization factor, which can be set as a constant, $k(x)$ is a kernel function, such as a Normal kernel as follows:

$$k_N(x) = C * \exp\left[-\frac{1}{2}\|x\|^2\right] \quad (2)$$

Then we need to initialize the new center location of the target in the current frame at the previous center location y_0 and compute the distribution of the target candidate at y_0 :

$$\hat{p}(y_0) = \{\hat{p}_u(y_0)\}_{u=1\dots m} \quad (3)$$

$$\hat{p}_u(y_0) = C_h \sum_{b(x'_i)=u} k\left[\left\|\frac{y_0 - x'_i}{h}\right\|^2\right] \quad (4)$$

where x'_i is the normalized pixel value in the target candidate which is defined to have the center at y_0 in the current frame, h is the bandwidth of the candidate area, and C_h is a normalization factor, which can be set as a constant.

To compare between the target model and the target candidate, the similarity function, $\rho(\hat{p}(y), \hat{q})$, is based on the Bhattacharyya Coefficient as follows:

$$\begin{aligned} BC1 = \rho(\hat{p}(y_0), \hat{q}_u) &= \sum_{u=1}^m \sqrt{\hat{p}_u(y_0)\hat{q}_u} \\ &= \sum_{u=1}^m \hat{p}_u(y_0) \sqrt{\frac{\hat{q}_u}{\hat{p}_u(y_0)}} \end{aligned} \quad (5)$$

Step 2: Derive the weights $\{w_i\}_{i=1\dots n_h}$, as follows:

$$w_i = \sum_{u=1}^m \sqrt{\frac{\hat{q}_u}{\hat{p}_u(y_0)}} ; i = 1 \dots n_h \quad (6)$$

Step 3: The mean shift vector computes the new location y_1 , by calculating the minimum distance between the target model and the target candidate. The current position y_0 will be moved to the new location y_1 . Therefore, the new location of the target candidate can be derived as follows:

$$y_1 = \frac{\sum_{i=1}^{n_h} x'_i w_i g\left(\left\|\frac{y_0 - x'_i}{h}\right\|^2\right)}{\sum_{i=1}^{n_h} w_i g\left(\left\|\frac{y_0 - x'_i}{h}\right\|^2\right)} \quad (7)$$

where $g(x) = -k(x)$

After updating the new center target location at y_1 , the distribution of target candidate at y_1 is computed as follows:

$$\hat{p}(y_1) = \{\hat{p}_u(y_1)\}_{u=1\dots m} \quad (8)$$

Then the second Bhattacharyya coefficient or the similarity function between the target model and the target candidate of the new location is evaluated.

$$BC2 = \rho(\hat{p}(y_1), \hat{q}_u) = \sum_{u=1}^m \sqrt{\hat{p}_u(y_1)\hat{q}_u} \quad (9)$$

Step 4: This process will iterate until $BC2 > BC1$; however, if $BC2 < BC1$, then the new center target y_1 will be updated as follows:

$$\begin{aligned} & \text{While}\{BC1 < BC2\} \\ & \text{Do}\left\{y_1 = \frac{1}{2}(y_0 + y_1)\right\} \end{aligned}$$

Step 5: This process will check the condition to terminate the algorithm based on the predicted threshold value ϵ . The threshold value is defined to be the minimum distance between the target model and the target candidate. The condition of this step will be computed as follows:

$$\begin{aligned} & \text{if}(\|y_1 - y_0\| < \epsilon) \\ & \quad \text{Stop the process} \\ & \text{Else} \\ & \quad \text{Set } y_0 = y_1 \\ & \quad \text{Go to Step 1} \end{aligned}$$

B. The Kalman Filter Algorithm

The Kalman filter [8] is based on a set of mathematical equations which implements a predictor-corrector step to estimate the result. This filter is a tool for estimating the states of a linear system. The Kalman filter is a recursive process which is separated into two steps consisting of prediction and correction steps. The prediction step defines the time update equations and the correction step defines the measurement equations. The goal of the Kalman filter is to determine a posteriori state estimate \hat{x}_k .

The time update equations, responsible for projecting in time, consist of the current state and the priori estimate error covariance for the next time step as follows:

- 1) Project the priori state, \hat{x}_{k+1}^- , in the next time, $k+1$, as follows:

$$\hat{x}_{k+1}^- = A\hat{x}_k + Bu_k + w_k \quad (10)$$

where A is an $n \times n$ state transition matrix, \hat{x}_k is the $n \times 1$ state matrix in the previous time (frame) step, n is number of estimate values, B is an $n \times l$ optional control transition matrix, l is the number of the control values, u_k is a k -time optional control input which is an $l \times 1$ matrix, and w_k is the process noise which is an $n \times 1$ matrix or a constant.

- 2) Project the priori estimation error covariance, P_{k+1}^- , in the next time, $k+1$, as follows:

$$P_{k+1}^- = AP_k A^T + Q \quad (11)$$

where $Q = [w_k w_k^T]$ is the process noise covariance which is an $n \times n$ matrix, and P_k is the posteriori estimate error covariance.

The measurement equations use the actual measurement z_k to update the state object consisting three processes as follows:

- 1) Compute the Kalman gain, K as follows:

$$K = P_k^- H^T (H P_k^- H^T + R)^{-1} \quad (12)$$

where K is called the ‘‘Kalman gain’’ which is an $n \times m$ matrix, P_k^- is a priori estimate error covariance, H is an $m \times n$ measurement transition matrix, $R = [v_k v_k^T]$ is an $n \times n$ measurement noise covariance matrix.

- 2) Update estimate state as follows:

$$\hat{x}_k = \hat{x}_k^- + K(z_k - H\hat{x}_k^-) \quad (13)$$

where \hat{x}_k^- is an $n \times 1$ priori state estimate matrix, z_k is an $m \times 1$ actual measurement matrix, which can be written as follows:

$$z_k = Hx_k + v_k \quad (13)$$

where v_k in the measurement noise which is an $m \times 1$ matrix or a constant, x_k is the result of the Target Representation and Localization part which is an $m \times 1$ matrix.

- 3) Update the posteriori estimation error covariance, P_k in the current time k , as follows:

$$P_k = (I - KH)P_k^- \quad (15)$$

III. ADAPTIVE MEAN-SHIFT KALMAN TRACKING

In this research, we propose the adaptive mean-shift Kalman tracking based on the mean-shift algorithm combined with the Kalman filter to track the object of interest object. Unlike the traditional mean-shift algorithm the proposed mean-shift algorithm can vary different ROI sizes of the target candidate. The overall process of adaptive mean-shift Kalman tracking summarized in Fig. 2 is explained below.

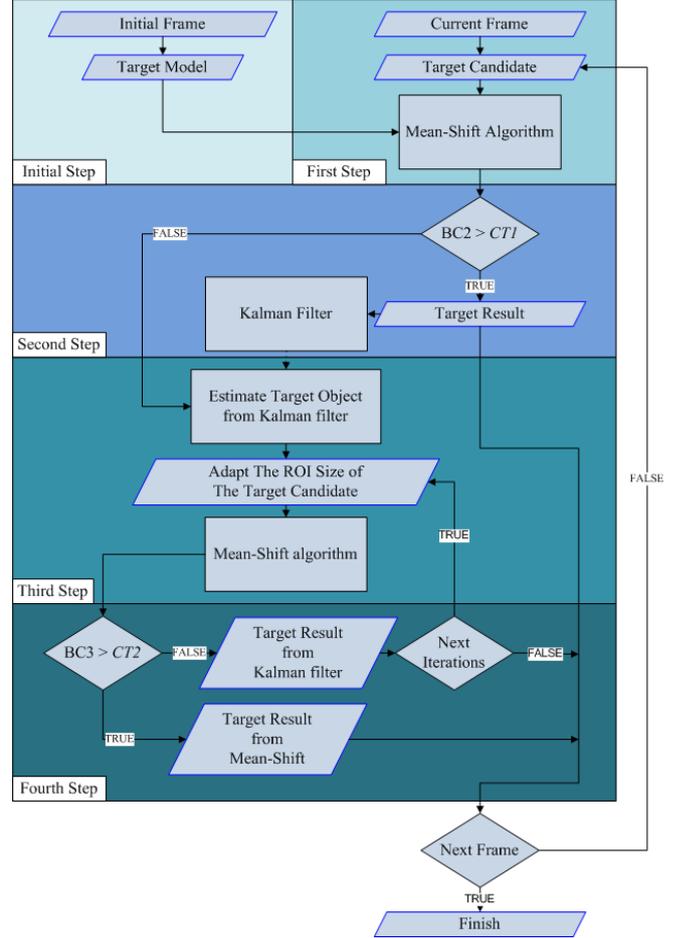


Fig.2. The overall process of the adaptive Mean-Shift Kalman tracking

Initial Step: Selection of target model

In the first frame, the user selects the target model which is the predefined ROI. From the target model, we will compute the initial state of the Kalman filter.

First Step: Mean-Shift algorithm process

In the next frame, the algorithm defines the target candidate which is at the same center location as the target model. However, the size of the target candidate is larger than the size of the target model which is the size of the region, h . The target model and the target candidate will be computed to get the second Bhattacharyya coefficient ($BC2$) in the mean-shift algorithm.

Second Step: Similarity comparison from the mean-shift algorithm

First, we will define the first similarity threshold value ($CT1$) to determine the tracking criteria. This value will be compared with $BC2$. If $BC2$ is more than $CT1$, then we will

update the Kalman filter by using the result of the mean-shift algorithm in the first step. Thus, the target result in this case will be acquired from the mean-shift algorithm, and the next sequence frame can be proceeded. However, if $BC2$ is less than $CT1$, then go to the third step.

Third Step: Estimation of the Kalman filter

In this step, the estimate state of the Kalman filter will feed back to the adaptive mean-shift algorithm. If the tracking result is not in the predefined similarity threshold value, the algorithm will increase the target candidate up to twice the size of the current ROI. In addition, this target candidate will define the new location from the current state of the Kalman filter. Hence, the third Bhattacharyya coefficient ($BC3$) is computed.

Fourth Step: Similarity comparison between the Kalman filter and the adaptive mean-shift algorithm

We compare the second similarity threshold value ($CT2$) with the $BC3$ from the third step. If $BC3$ is greater than $CT2$, then we will use the result from the adaptive mean-shift algorithm. If $BC3$ is smaller than $CT2$, then we will use the estimate state of the Kalman filter, increase the target candidate ROI and go back to the third step. This process is repeated until the maximum target candidate size is met. Otherwise, the target result will be acquired from the adaptive mean-shift algorithm, and the next sequence frame can be proceeded.

IV. EXPERIMENTAL RESULTS

From the object tracking system, we have proposed the adaptive mean-shift kalman algorithm to locate the target object which is the laparoscopic instrument. To illustrate the performance of the proposed adaptive mean-shift kalman algorithm, we tested on two situations which consist of simulated videos, and real laparoscopic videos which were derive from a sub-process of appendectomy [12].

A. Simulated Video Situations

To test the proposed algorithm, we used the simulated videos for different situations. From the simulated videos in Fig. 3, the red background represents the color of the internal body in laparoscopic surgery. The yellow ball represents an obstacle. The tip of the real laparoscopic instrument represents the target object, which is shown in the white rectangle. The green rectangle in the tracking process is the result of the tracking algorithm. In the first frame, we need to initialize the target model for the proposed algorithm as shown in Fig. 3.



Fig.3. The target model with an initial value in the white rectangle

Simulated Experiment #1: Change the shape of the target model and move in front of the obstacle

The goal of this experiment is to track the target object when changing the shape of the target and moving the target in front of the obstacle. This experiment was computed on a video with 196 frames covering about 20 seconds. Thus, the average of frame rate is 9.8 frames per second. Fig. 6 shows fifteen sample frames at different times. This experiment shows that the proposed algorithm can track the target object with correct locations.

Simulated Experiment #2: Resize and change the shape of the target

The goal of this experiment is to track the target object when changing the size as well as the shape of the target. This experiment was computed on a video with 240 frames covering about 25 seconds. The tracking results in some frames are shown in Fig. 7. The proposed algorithm can track the target object in all frames.

Simulated Experiment #3: Move behind the obstacle

The goal of this experiment is to track the target object when the object moves behind the obstacle. This experiment was computed on a video with 140 frames covering about 14 seconds. Some tracking results are shown in Fig. 8. From 4-12 seconds, the whole target object was hiding behind the obstacle. After that, the target appeared in the scene again and the tracking process can track the target correctly. This is due to the estimation process in the Kalman filter to improve tracking performance. Without the Kalman filter, we were not able to track the target object correctly after it disappeared from the scene.

B. Real Laparoscopic Video Situations

The real situation videos are derived from the real laparoscopic surgery in the operating room [12]. Real laparoscopic surgery is more complicated than the previous simulated situations, because it combines different situations at the same time.

Real Laparoscopic Surgery Experiment #1

In this experiment, we tested our proposed algorithm on the real video of laparoscopic surgery, which is a sub-process of appendectomy [12]. This process cut a cord using a serrated scissors instrument which is a laparoscopic instrument. This experiment video has 132 frames which cover about 16 seconds. Thus, the average of frame rate is 8.25 frames per second. We initialized the target model, i.e., the tip of the laparoscopic instrument, with the blue rectangle as shown in Fig. 4.



Fig.4. The target model with an initial target in the blue rectangle for tracking in real laparoscopic surgery experiment 1

Fig. 9 shows the result of adaptive mean-shift Kalman algorithm representing in the orange rectangle in the each image. This figure represents the examples of fifteen frames. Here, the target object not only changes its size and shape, but also hides behind another laparoscopic instrument. The overall experiment results show that the adaptive mean-shift Kalman algorithm can locate the tip of laparoscopic instrument which is in real laparoscopic surgery which covers the overall situations in this video.

Real Laparoscopic Surgery Experiment #2

This video coagulates a tissue by an electrode instrument. This experiment video has 205 frames which cover about 28 seconds. Thus, the average of frame rate is 7.32 frames per second. At the first frame, we initialized the target model, as shown in Fig. 5.



Fig.5. The target model with an initial target in the blue rectangle for tracking in real laparoscopic surgery experiment 2

The overall results of this case will show in Fig. 10. The challenge problem of this experiment is the smoke which has occurred from the electrode instrument, especially during 10 to 40 seconds. The proposed algorithm can track the target object in the near correct location. After that, the target object has changed its size, shape, and background together with little smoke. However, the proposed algorithm can still track the correct location of the target. The overall experiment results show the adaptive mean-shift Kalman algorithm can locate the tip of the laparoscopic instrument even when it is covered by smoke.

V. CONCLUSION

In this paper, we have proposed the adaptive mean-shift Kalman tracking based on the mean-shift algorithm combined with the Kalman filter. The ROI size of the target candidate at each frame can be adjusted to increase the chance of tracking. The experimental results of simulated situations show that the proposed algorithm can locate the target object correctly even when the size and the shape of the target have been changed. In addition, when the target is hiding behind some obstacles, this algorithm can still track the target object after it comes out. Likewise, the experimental results of real laparoscopic surgery show that the proposed algorithm can locate the target object even where the target object hides behind another surgical instrument, and when smokes cover the surgical site. In addition, this algorithm can locate the target object that does not depend on the type of laparoscopic instruments. Thus, this proposed algorithm will be suitable for locating the tip of the laparoscopic instrument as well as guiding the path of our conceptual robot in real laparoscopic surgery.

ACKNOWLEDGMENT

The authors would like to thank Thailand Graduate Institute of Science and Technology (TGIST) for their financial support.

REFERENCES

- [1] R. H. Taylor, J. Funda, L. Joskowicz, A. D. Kalvin, S. H. Gomory, A. P. Guezic, and L. M. G. Brown, "An overview of computer-integrated surgery at the IBM Thomas J. Watson Research Center," *IBM J. Research and Development*, vol. 40, pp. 163-183, 1996.
- [2] K. Omote, H. Feussner, A. Ungeheuer, K. Arbter, A. Guo-Qing Wei, "Self-guided robotic camera control for laparoscopic surgery compared with human camera control," *The American Journal of Surgery*, vol. 177, pp. 21-24, 1999.
- [3] A. Casals, J. Amat, E. Laporte, "Automatic guidance of an assistant robot in laparoscopic surgery," *International Conf. on Robotics and Automation*, vol. 1, pp. 895-900, 1996.
- [4] C. Lee, Y. F. Wang, Uecker, Y. Wang, "Image analysis for automated tracking in robot-assisted endoscopic surgery," *Proceedings of the 12th IAPR International Conference*, vol. 1, pp. 88-92, 1994.
- [5] A. Guo-Qing Wei, G. Hirzinger, "Real-time visual servoing for laparoscopic surgery. Controlling robot motion with color image segmentation," *IEEE Engineering in Medicine and Biology*, vol. 16, pp. 40-45, 1997.
- [6] V. Sa-Ing, S. S. Thongvigitmanee, C. Wilarusmee, J. Suthakorn, "Adaptive Mean-Shift Kalman Tracking for Laparoscopic Surgery," *the 4th IEEE International Conference on Computer Science and Information Technology (IEEE ICCSIT 2011)*, June 10-12, 2011, pp. 204-208.
- [7] D. Comaniciu, P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence* vol. 25, pp. 603 – 619, 2002.
- [8] C. Masreliez, and R. Martin, "Robust bayesian estimation for the linear model and robustifying the Kalman filter," *IEEE Transactions on Automatic Control*, vol. 22, pp. 361 – 371, 2003.
- [9] V. Sa-Ing, S. Sothivirat, C. Wilarusamee, J. Suthakorn, "Design of A New Laparoscopic-Holder Assistant Robot," *the 3rd International Symposium on Biomedical Engineering (ISBME 2008)*, Bangkok, Thailand, pp. 278 – 281, 2008
- [10] V. Sa-Ing, S. S. Thongvigitmanee, C. Wilarusmee, J. Suthakorn, "A New Laparoscopic-Holder Assisting Robot and Workplace Analysis," *the 7th Asian Conference on Computer-Aided Surgery (ACCAS 2001)*, Bangkok, Thailand, 2011.
- [11] A. Yilmaz, O. Javed and M. Shah, "Object Tracking: A Survey," *ACM Computing Surveys*, vol. 38, pp. 1 - 45, 2006.
- [12] Mishra, R.K., Free Surgical Laparoscopic Videos, in <http://www.laparoscopyhospital.com/DOWNLOADS.HTM>, lap_appendectomy.wmv, Editor. 2008.



laparoscopic surgery, and medical robotics.

Vera Sa-Ing was born in Trat, Thailand on September 30, 1983. He received the B.S. degree in Computer Engineering from Suranaree University of Technology, Thailand in 2005 and the M.S. degrees in Biomedical Engineering from the Mahidol University, THAILAND in 2011. From 2005 to 2007 he was a research assistant working at the National Electronics and Computer Technology Center (NECTEC). His research interests include image processing, tracking algorithm, medical imaging,



Technologist Award from the Foundation for the Promotion of Science and Technology under the Patronage of H.M. the King.

Saowapak S. Thongvigitmanee received the B.S. degree in Electrical Engineering from Cornell University, USA in 1997, and the M.S. and Ph.D. degrees in Electrical Engineering from the University of Michigan, Ann Arbor, USA in 1998 and 2003, respectively. Since 2003, she has been working at the National Electronics and Computer Technology Center (NECTEC). Her research interests include image processing, signal processing, and medical imaging. In 2008, she won Thailand Yong



Chumpon Wilasrusmee received the M.D. degree (First Class Honor) from Ramathibodi Hospital, Mahidol University, Thailand in 1992 and Thai Board of General Surgery in 1998. From 1999-2004, he had been a research fellow in the Transplantation Services, Department of Surgery, SUNY Upstate Medical University, Syracuse, NY. Since 2008, Dr.Chumpon became a Professor of Surgery, Faculty of Medicine, Ramathibodi Hospital.



Jackrit Suthakorn received his Bachelor's degree in Mechanical Engineering from the Mahidol University in 1995. He received his Master degree in Automatic Controls from the Michigan Technological University, and Ph.D. in Robotics from the Johns Hopkins University, Baltimore, Maryland, in 1998 and 2003, respectively. Presently he has resumed his faculty position at the Mahidol University, where he is now the Director of the Center for Biomedical and Robotics Technology (BART LAB), His interest includes Robot-Assisted Surgery, Surgical Navigations, Self-Replicating Robots, Binary Hyper-Redundant Robotic Manipulators and MEMS.

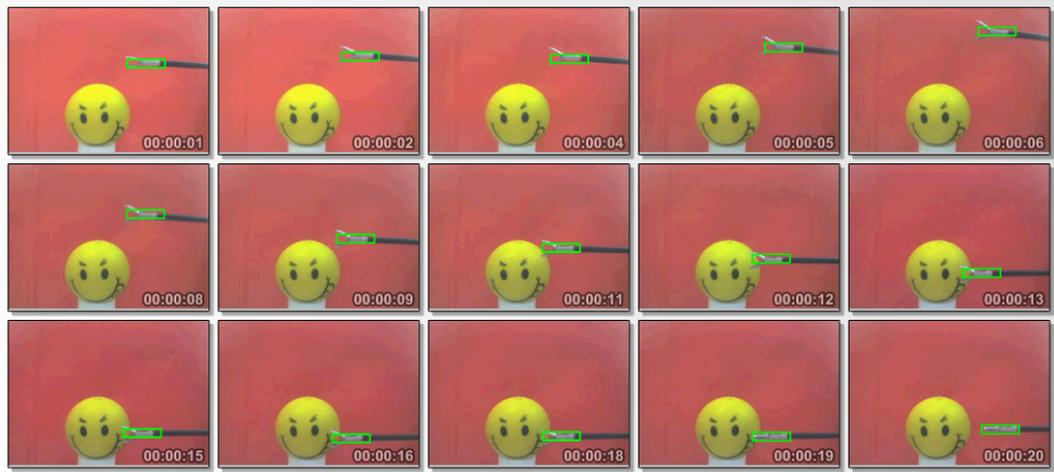


Fig. 6. Sample results in Simulated Experiment #1 using the proposed algorithm



Fig. 7. Sample results in Simulated Experiment #2 using the proposed algorithm

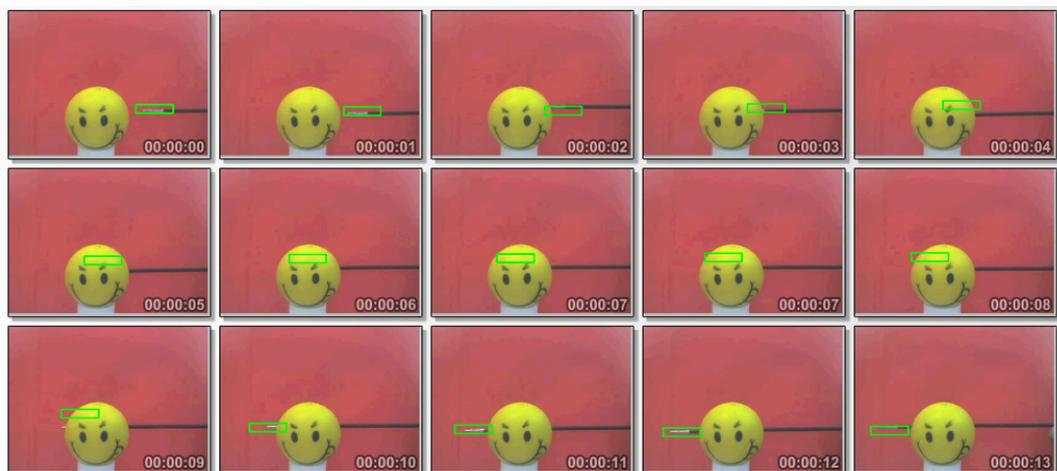


Fig. 8. Sample results in Simulated Experiment #3 using the proposed algorithm

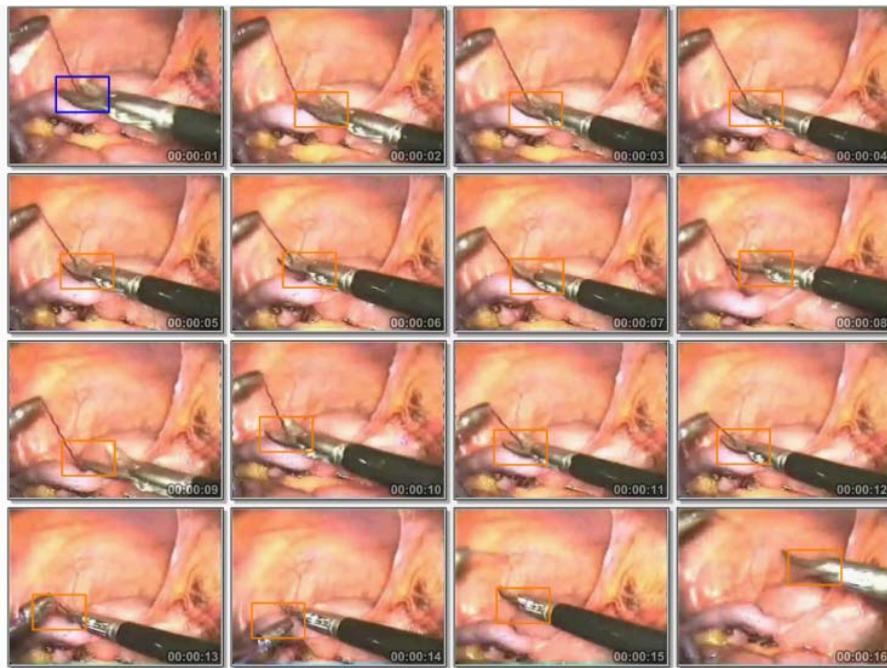


Fig. 9. Sample results in Real Laparoscopic Surgery Experiment #1 using the proposed algorithm

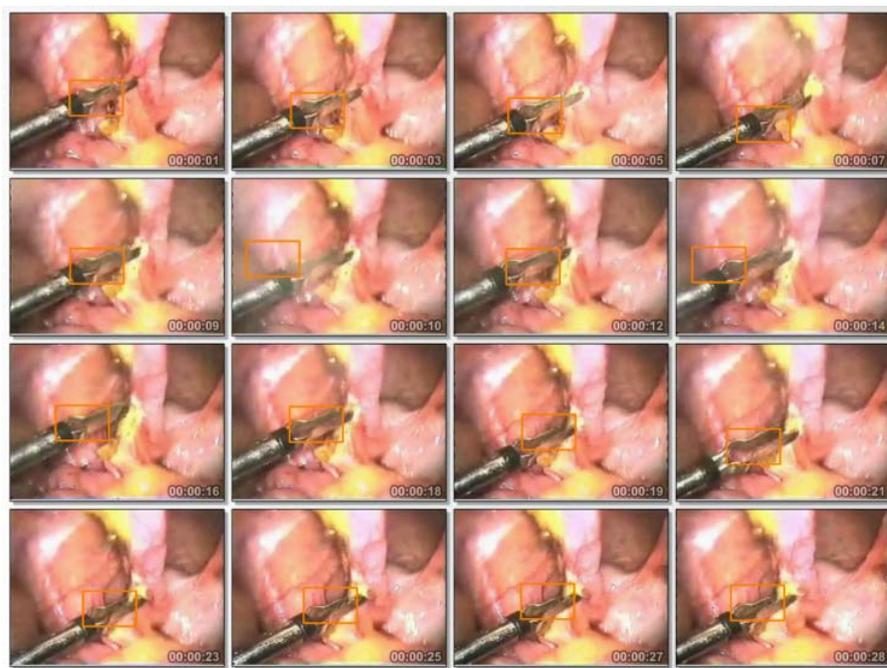


Fig. 10. Sample results in Real Laparoscopic Surgery Experiment #2 using the proposed algorithm