Journal of Imaging Science and Technology® 53(2): 020506–020506-10, 2009.

Tracking Objects with Radical Color Changes Using a Modified Mean Shift Algorithm

Inteck Whoang and Kwang Nam Choi
School of Computer Science and Engineering, Chung-Ang University, Seoul 156-756, Korea
E-mail: knchoi@cau.ac.kr

Samuel Henry Chang
School of Electrical and Electronics Engineering, Chung-Ang University, Seoul 156-756, Korea

Abstract. *This paper presents a new method for tracking moving objects with radical color changes using modified mean shift. The mean shift algorithm seeks the highest density value using a mean shift vector obtained from the density gradient. Density presents various forms of object information such as color and intensity depending on the application. Conventional color-based mean shift methods show good results when tracking nonrigid moving objects. However, they do not provide accurate results when the initial color distribution of the object disappears. In our method, color distribution is used to represent the objects. The mean shift algorithm is first used to derive an object candidate by estimating the maximum increase in density direction from its current position. Next, the color variation of the object is calculated and compared with a specific threshold value. When the color variation of the object exceeds this threshold value, the initial color of the search window is updated. The objective of our method is to provide for robust real-time object tracking with large color variation in the object whose color changes during motion. The implementation of the new algorithm shows effective tracking results with complete object color changing from time to time. Validation of our approach is illustrated by comparison of experimental results obtained using the methods described above. © 2009 Society for Imaging Science and Technology.*
[DOI: 10.2352/J.ImagingSci.Technol.2009.53.2.020506]

INTRODUCTION

The subject of real-time object tracking can be described as a correspondence problem involved with locating an object from frame to frame in a video image sequence. Normally, the time interval between two successive frames is small and constant. Therefore, the interframe changes are limited, and the correspondence can be created based on features extracted from the image frames. However, object feature extraction can be very sensitive to variations in illumination, viewpoint, scale occlusion of the objects, and the complexity of its background. The appearance of the object can be changed under inconstant working conditions. To keep track of an object when its features change, the most important issue is to enhance and update the object's feature information with regard to the noisy and dynamic background. A template approach was developed to prevent the "drifting" problem by using a heuristic algorithm.[1] The method was also extended to convert a generic Active Appearance Model to a specific model. Bayesian bootstrap and particle filters were developed to ensure robust tracking of objects with occlusion, rotation, and scale variation.[2,3] Robustness was obtained by updating the weighted average between its current and new histogram. An advanced tracking system was constructed based on a Gaussian Mixture Model (GMM); this model requires a fixed number of Gaussian components, including a stable component.[4] Gaussian Mixture Models based on histogram information using Expectation Maximization (EM) algorithms work well in situations where the observation data are incomplete.[5] They start with an initial estimate and iteratively update its estimation to increase the likelihood of the observed data, and find the expected value of the complete-data log-likelihood with respect to the unknown data, given the observed data and the current parameter estimates. In order to use the EM algorithm, we need to ensure the number of components in GMM and avoid adapting extra modes in the underlying density function or frequent mode switching.

Most object tracking systems require robustness toward arbitrary noise and tolerance to large variations in the observation data. We can ensure the requirements of the target through prior knowledge and proper processing. However, it is especially difficult to track an object in dynamic and complex circumstances. For example, in color-based object tracking, the object features are represented through color distributions. The object's feature appearance can change from time to time due to lack of color constancy capabilities. Furthermore, the object's features can disappear when its color is changed completely. Therefore, color-based algorithms with a color constancy assumption have been suggested, but their performance has been inadequate thus far. In applications such as tracking human motion, the color of the human object can change when, for example, clothes are changed during the tracking time. In a facial tracking applications, one's face color distribution can change radically when one shakes or raises his or her head because the orientation of the face with respect to the camera is changed. The initial face color can disappear completely when one turns over and the tracking camera captures the back of the head, as shown in experiment at the end of this paper.

Received Nov. 7, 2007; accepted for publication Jan. 23, 2009; published online Mar. 10, 2009.
1062-3701/2009/53(2)/020506/10/$20.00.

Mean shift is a nonparametric algorithm that shifts each sample to the model it belongs to statistically.[6] It can be generated like clustering algorithms.[7] It was originally developed for data clustering and was later extended to delineate arbitrarily shaped clusters for face tracking.[8] Color-based mean shift algorithms have been used to track human motion. Mean shift analysis was used to derive the target candidate that has the most similarities to a given target model, and thus to track objects with changing scale. It is modified to deal with dynamically changing color distribution and to find the mode of color distribution within a video scene.[9–12] An updated mean shift-based model technique was used to improve a robust and speedy human tracking system.[13] A color-based mean shift tracker and the sum of square differences were used in the Kalman filter framework to track far reaching objects ranging from surveillance to smart rooms.[14] Mean shift clustering was used to detect moving shadows and to recover significant image features.[15,16]

Examples of tracking of an object where color changes radically were shown in some color-based mean shift object tracking algorithms.[8,9] However, cases with complete color change have not been explored and examples where the initial color changed completely were excluded. In this paper, we propose a modified mean shift algorithm with robust performance in tracking an object with radical and complete color change. The term "radical and complete" color changes can be defined as a complete color change when the original color of the tracking object disappears in its entirety and is replaced by a different color.

Our approach is as follows. First, the color distribution of an object is initialized. Second, the object position is determined using color distribution in each frame using mean shift analysis. Third, the color distribution of the object is updated using the object position. In this paper, a modified mean shift algorithm is used iteratively to update the density function in order to ensure the robustness of the track with respect to a radical change of color distribution.

RELATED WORK

A mean shift algorithm is a simple iterative procedure that shifts each data point to the average of data points in its neighborhood; it is a general algorithm that may be used for data clustering, analysis, and computer vision.[17] The basic idea of the mean shift algorithm is to reach the nearest maximum density peak from the gradient of a probability distribution. The search is based on the kernel density estimation that converges by moving from the current location to the location with the highest density iteratively. In particular, a mean shift tracking algorithm based on color is a real-time algorithm that seeks to maximize the correlation between two statistical distributions; i.e.., to represent the similarity between the two distributions. Statistical distributions can be built using the characteristics of a tracked object discriminated from its surrounding environments.

The mean shift algorithm has several advantages; it does not require *a priori* assumptions for global density because it only uses gradients from neighbor distributions, while iterative processing makes it simple. Besides, when mean shift is applied to object tracking, mean shift that exploits only color information without any shape features is very useful for tracking a nonrigid object. Mean shift also has good performance when tracking rigid objects with constant colors. However, mean shift cannot solve the traditional local minimum problem as completely as almost all other clustering algorithms do. Moreover, mean shift cannot track its target correctly when its color changes.

The latest work using a similar methodology is the approach taken by Stern and Efros.[18] They track the human face under varying illuminations and color-changing targets by switching color spaces, using mean shift and a Color Distribution Model (CDM) incorporating a probability distribution function (PDF). The CDM/PDF/color histogram is updated with a damping factor in each frame. Multiple CDM creation methods were tested. The significant difference between their approach and the present approach is as follows. First, while both approaches use CDM, they update the CDM/PDF/color histogram with a damping factor in each frame, while we only update the search window when the color variation is larger than a predetermined threshold value. Experimental results below show that the their approach fails when a target is moving close to a background with a color similar to the target, while ours tracks the target successfully under the same condition because the absence of a color variation threshold which can induce incorrect responses. Second, we use a single search window while their approach requires two: an internal window and an external window. The internal rectangular window contains the object to be tracked. The external window is larger than internal window and is used to estimate the best color space. The computational time of the CDM in this approach is faster because only a single search window is used. Third, we use only a single color space model while they switched from a previous selection to the best model after calculating and comparing with each color space model.

THE PROPOSED ALGORITHM

Conventional color-based tracking algorithms using mean shift do not use any information regarding the shape of the target. They also assume that the target object maintains its initial colors, though it changes its silhouette during the motion. Before tracking an object, the algorithm initializes the color distribution of the object with predefined measurements. This initial information is then used to track the location of the target during motion. It is possible to track a nonrigid object with a fixed color distribution; however, when the color distribution of the target object changes, the algorithm may yield an incorrect location of the target because it only estimates the target location based on its initial color information. It is also difficult to separate the nonrigid object from its background because the target and the background could mix. Furthermore, nonrigid objects, such as a walking human or a running animal, can change their outlines when moving. It is very difficult to predict a target's shape in next frame based on current information. In this

Figure 1. An example of the initial image frame with an experiment showing a target cup with an initial color of yellow.

case, the best solution is to update the color distribution of the target whenever its color changes. The basic idea behind the proposed algorithm is to detect the changes of color distribution of the target initially and to update the target's information when the detection response is positive. The algorithm is illustrated step by step with some examples in the following subsections.

Target Object Color Initialization

The initialization of the object's color is given manually and in this respect there is no difference between the conventional mean shift algorithm and the proposed algorithm. However, to show the differences between the two algorithms, the target in this example is selected as a paper cup with two distinct colors: solid yellow on one side and violet on the other. The location and the color of the target are changed from frame to frame when the cup is moving. The initial position and the initial color of the cup are shown in Figure 1.

A search window is shown as a rectangle over the image frame used to model the target by a set of pixels within the window. It has a specific predefined size that is centered at the position of the target and indicates a sample data set to model the target. It is used to obtain the density of its neighborhood and to detect color variation on the target. The size and location of the search window present the appearance of the target over the image frame. The search window is defined by the object detection application. The size of the search window is determined by the target object and can be defined automatically after its initial color and position is given. In order to focus on the search window update, the size of the search window is manually designated in this paper.

There are three histograms of RGB color used to model the moving target in this paper. Color histograms show the color distributions for all color components. The entire color distribution can be maintained constantly in most

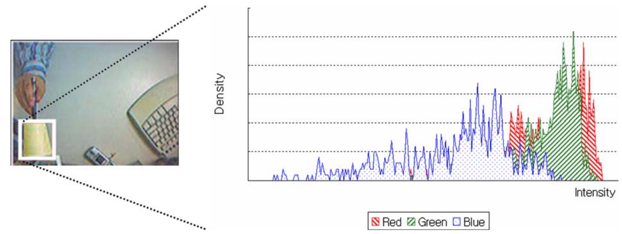

Figure 2. The search window is initialized to contain the targeted cup as shown by the white rectangle in the image on the left of the figure. The RGB histogram of the target object within the search window is shown on the right of the figure.

Figure 3. This image shows that the location and color of the cup changed when the cup moves.

cases even if the target changes its color or shape. Therefore, the color histograms are appropriate to measure the color similarity of the objects between two image frames. The RGB histograms of the object are obtained within the search window. The search window is initialized to a predefined size. The initial position of the target is selected as the center of the search window. The plot in Figure 2 shows the histograms of RGB components in the search window from the initial frame. The horizontal axis represents the intensity of each color component, while the vertical axis indicates the density of each level.

After the color of the object is initialized, the histograms of the RGB components in the search window from the initial frame is then used to check for any location and color change from one frame to another.

Figure 3 shows an example of such an image frame. The location and color changes to the target can be calculated by comparing the current and the previous frames (the second and first frames in this example). In the image frame shown in Fig. 3, the location of the cup has changed from bottom left of the initial image in Fig. 1 to bottom middle of the image. The color of the cup has also changed from its initial yellow to yellow and violet.

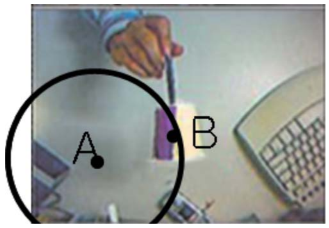

Figure 4. The locations of the target during the motion are estimated by the density of the neighborhood of the search window. "A" shows the previous target location from Fig. 1, while "B" shows the current target location.

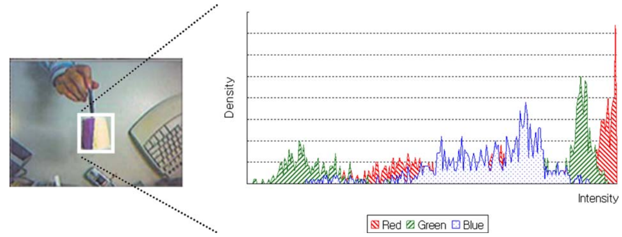

Figure 5. Search window has been updated and is shown as the white rectangle in the image on the left of the figure, and the updated color histogram of the search window is shown on the right of the figure.

Target Location Estimation

The color histograms of each neighbor for all orientation can be obtained by the identical processing as shown in Fig. 3. These histograms are used to obtain the PDF. The PDF is defined as

$$\rho[\widehat{x_0},\hat{q}] = \sum_{n=1}^{m} \sqrt{f_n(\widehat{x_0}) f_n(\hat{q})}, \tag{1}$$

where x_0 represents the initial target object, q is a neighbor set, f_n represents the frequency of the three histograms at nth level, and m is the number of levels in f. The PDF reflects the probability of target position in the next frame. The location with the highest density indicates the highest probability of target position in the next image frame. The location of the highest density can be obtained from the PDFs of all its neighbors.

Figure 4 illustrates these processes. Point "A" indicates the center of the search window of target from its previous image frame. The radius of the circle with origin "A" represents the range of the search area. The radius length represents a tradeoff between computation time and accuracy. It determines how many neighbors need to be computed. A longer radius needs a longer computational time, and a smaller tracking error can be reached. In this paper the maximum length of the radius is defined as the distance of the target between two frames. After the histograms of all the neighbors inside the circle are estimated, the PDF is calculated using Eq. (1). The highest density location is selected by comparing of two image frames. As one can see from Fig. 4, although the targeted object changed its position and color in the current frame, the yellow color partially remains. As shown in Fig. 4, point "B" has the highest PDF. Therefore, it is considered as the target position for the next frame.

Target Color Variation Estimation and Search Window Update

When the object has only translation or small color variation with little influence on tracking, then it can be tracked accurately by mean shift. However, in some cases, an object may have larger color variations. For instance if the target is the face of a person with varied hair color and the target turns his head so that his face disappears completely, the mean shift will not perform properly. Conventional mean shift assumes that the target keeps its initial colors when

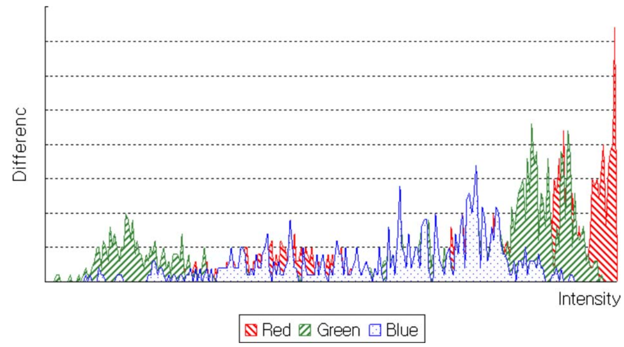

Figure 6. The histogram differences between initial color distributions within the search window in Fig. 2 and the current color distributions within the search window in Fig. 5.

tracking. Therefore, the search window needs to be updated with the large color variations of the current target. Our proposed algorithm is developed to track the target in such circumstances. In this algorithm, color variation of the target is calculated as the first step to update the search window. The value of color variation is computed using following equation:

$$\nabla C = \sum_{n=1}^{m} |f_n(\widehat{x_0}) - f_n(\widehat{q_h})|, \tag{2}$$

where q_h is a neighbor with the highest density, ∇C represents the value of the color variation, and the other parameters are the same as those used in Eq. (1). When the color variation exceeds a specific threshold, the search window is updated as the highest density location. Figure 5 shows three histograms for the updated search window as the highest density location. The difference between histograms in this figure and those in the Fig. 2 reflect the change in the target color from yellow to violet.

The difference in intensity for all the color components between the initial color distributions and the subsequent color distributions from the next frame within the search window is shown in Figure 6. There are large variations between the two overall intensities. The color changes on the target can then be detected through the sum of ∇C for all the color components in the plot according to Eq. (2). The target location is then updated by comparing the position between the previous frame and the highest density location in the current frame. An accurate position can be calculated when the value of the difference converges by repeating the process by computation according to Eq. (1). The procedure of the algorithm is explained below.

The Proposed Algorithm

The process of the proposed algorithm can be implemented by following these steps:

Step 1: Initialize the search window using the density function of the target object, window size and color bin size.

Step 2: Obtain the next image frame $i \leftarrow i+1$; (i: time sequence).

Step 3: Calculate the target location using mean shift.

Step 4: Update search window as current density function by obtaining the highest density location if necessary.

Step 5: Repeat step 2 until the difference value converges.

The parameters to perform the proposed algorithm, such as the size of the search window and the size of the histogram, are initialized in *Step 1*. The target model is then estimated using the specific search window. In this paper, the RGB histograms are used as the target model. In next frame, conventional mean shift is used to locate the target. The color variation is then calculated using the difference in the color variation between the search window from the previous frame and highest density location in the current frame. When the difference in the color variation exceeds a certain threshold value, the search window is updated as the highest density location in *Step 4*. The update to the search window guarantees that this algorithm will track the target correctly even though the target changes both its color and location.

The appropriate threshold is independent of the target size because the color histogram has already been normalized when the threshold is determined. The value of threshold is determined in practice by the proportion of noise to the target model defined as color histogram in the search window. The noise can be background noise, or any unexpected color noise. Unexpected color noises can include color variation noise from the frame to frame. When the color component histograms of the target are computed using the algorithm, the color variation of the target calculated may not be accurate because of such noise. The purpose of using the color variation threshold is to ensure that the algorithm is stable with respect to unexpected noise source. The threshold value is designated as a value larger than the possible color variation estimation error. When there is no information on the color variation of the target in the various video sequences, a suitable threshold value may not be determined. In this paper, the color variation threshold value is predefined by estimating the noise rate in the search window. The best threshold for various video sequences can be measured using a conventional training algorithm although this may be done without any prior knowledge of the noise. With a newly defined threshold value, the search window will only update when the difference in the color variation is larger than the threshold value. It will not update if a color variation within the estimated margin of error occurs. The search window is updated when a larger change in color of the target is detected (see below).

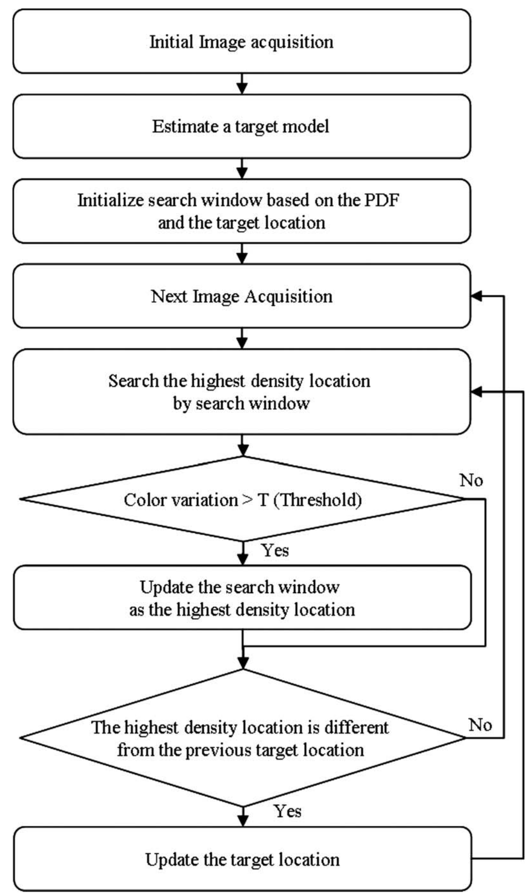

Figure 7. Flowchart of the proposed algorithm.

The steps of the algorithm are also illustrated in the flowchart shown in Figure 7. From the flowchart one can see that the step to detect the color variation is the key to this algorithm: the process of the steps before color variation detection is the same as in the conventional mean shift computation.

EXPERIMENTS AND COMPARISON

The proposed algorithm has been implemented and tested in experiments and compared with other methodologies. The experiments were undertaken in an actual setting and performed with human targets and artificial objects with interior illumination. Two different colors are used in the artificial object to test how effectively the algorithm tracks the motion of an object when it moves and changes its color. The colors are well separated from each other in the color space. A normal PC webcam with a resolution of 320×200 pixels with a capture speed of 15 frames per second is used. With the focus on the color variation, target is assumed to be at constant scale during tracking.

Three experiments were designed to test how effective this method tracks:

(1) moving object with radical color changes,
(2) moving object with complex real world background,

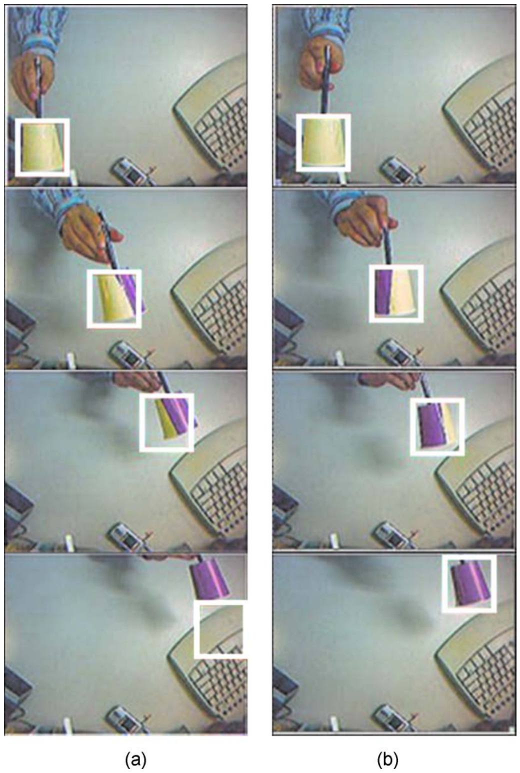

Figure 8. Experimental tracking results of an artificial cup object with radical color changes: (a) using the conventional mean shift analysis and (b) using the proposed modified mean shift algorithm.

(3) human motion with color changes.

The experimental results were also compared with outcome from a conventional color-based mean shift object tracking algorithm. The experiments and tests described below show that this algorithm successfully tracks moving objects with complete color changes, while the conventional algorithm lost its tracking when the object's color was changed completely.

Tracking Moving Object with Radical Color Changes

This experiment is designed to test how effectively this method tracks the moving object with radical color changes. An artificial object, a cup, was painted with two different colors—yellow and magenta—on both sides. The cup was designed to show its yellow side to the camera initially when it starts to move. Then, it moves across a table with computer keyboard and mouse and rotates itself so that the initial yellow side is slowly replaced by the magenta side of the cup until the yellow side has completely disappeared.

The tracking outcomes of both algorithms are shown in Figure 8. Fig. 8(a) shows a subset of the sample image frame using the conventional color based mean shift tracking algorithm, while Fig. 8(b) shows the sample image from the proposed algorithm. The sample images show a large color location shifting by temporal flow. Tracking results using the method shown in Fig. 8(b) successfully tracks the object when the cup changes its color from yellow to magenta. Tracking results using the conventional mean shift shown in Fig. 8(a) shows an inaccurate object positions during tracking and the algorithm lost track of the mug in last frame. This is because mean shift calculates the highest probability position based on its initial probability distribution with no knowledge in color variations. For those image frames containing the initial yellow color, mean shift method tracks the color and sets it at the center of the search window.

From this experiment, we may conclude that the conventional approach based on color information exploits the background information or shape data from the object. It lost track of the object at the end because it fails to update the initial matching color and thus uses inaccurate features. The proposed algorithm updates the initial color whenever the color of the tracked object is changed and is able to track the object even when it changes its color completely.

Tracking Moving Object with Complex Real World Background

This experiment is designed to test how effectively the method tracks the moving object with radical color changes against a complex background. The target is a mug with two different colors; one half of the mug is yellow while other is red. The target mug is designed to move with radical color changes. It will change its color and location concurrently and it changes color completely at the end. The moving background is a real world condition with a color distribution similar to that of the target. This experiment is performed under an identical environment.

Both tracking outcomes of our algorithm and the conventional mean shift method in this experiment are shown in Figure 9. Fig. 9(a) shows the experimental results from conventional mean shift method, while Fig. 9(b) shows the results from our algorithm. These results show that this algorithm tracks the target accurately while the target is moving with a large variation of colors and locations, and changes to a completely different color at the end. However, the results from the conventional mean shift tracks the target correctly during the target motion with initial yellow color distribution. The white rectangle indicates the estimated location of the target. However, since it tracks the object by color and the target switches to a towel with similar color at the end.

From this experiment we may conclude that the conventional approach based on color information exploits the background information or shape data from the object. It fails to track when another object with similar color to the initial object's color is located. The proposed algorithm can track the object even with color changes because it calculates the color variation iteratively on every frame. When the value of the variation exceeds a specific threshold, the search window is updated to the new target. The updated search window indicates the location of the new target accurately.

To inspect the color changes of the target during movement, identical sizes of each search window showing the tracking target using the proposed algorithm in the experiment are shown in Figure 10. From the image sequence, the target mug changes its color from its initial color of yellow to red. The target from Fig. 10(5) begins to change to red. The

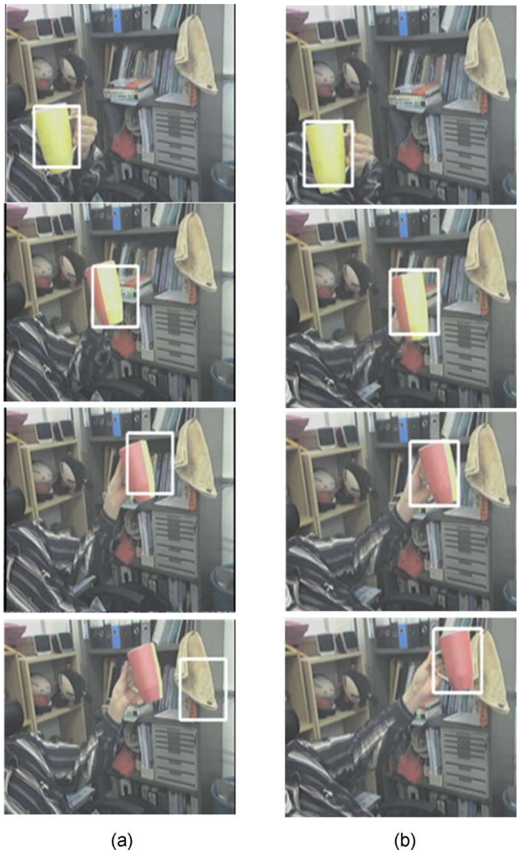

Figure 9. The experiment results of two tracking methods for a mug against a complex real world background: (a) results of conventional mean shift method and (b) results of our method.

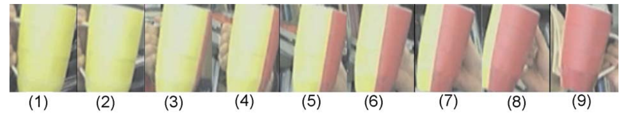

Figure 10. Sequence of the target object shown in Fig. 9. Each number indicates frame number. The target changes its color rapidly from yellow to red.

red color area over the target becomes larger than the area of the initial color of yellow as shown from Fig. 10(7) to Fig. 10(9) of the figure. Finally in the last frame shown in Fig. 10(9), its initial color disappears completely.

The estimation of color variation of the target during the experiment is shown Figure 11. The horizontal axis represents frame number with respect to time and vertical axis indicates the value of color variation calculated using Eq. (2). The frame number corresponds to the frame number of Fig. 10. The color variation is obtained by computing difference between initial target and subsequent targets. The initial target is frame number 1 corresponding to Fig. 10(1). As one can see, zero color variation of frame 1 in Fig. 11 represents the initial color of the search window of Fig. 10(1). A value of 25 000 calculated using Eq. (2) shows the maximum color variation between the initial yellow color and the current red color. The values of the color variation in each frame are used to monitor the object color changes and are used to determined whether the search window will be updated.

Position tracking accuracy is a major factor in evaluating an algorithm. The comparison of location tracking performance between conventional mean shift and the proposed algorithm is shown in Figure 12. In this figure, a location error is defined as a Euclidean distance between the target's true location and its measured location in pixels. The true locations of the target in each frame are defined manually. The measured locations in each frame are calculated from the center of search window. The location errors for both the mean shift algorithm and the proposed algorithm are plotted in Fig. 12. As one can see, both algorithms share similar performance from frame 1 to frame 4. However, a larger location error occurs from the mean shift algorithm in frame 5, where a larger color change of the target mug occurs. The location error from the mean shift algorithm reaches 120 pixels when it lost track of the target completely. The comparison of the tracking location errors between the mean shift algorithm and the new algorithm shows that the proposed algorithm is robust in position tracking.

Tracking Human Motion with Color Changes

This experiment is designed to test how effectively the method tracks human motion with color changes. The tracking object selected is a human head. A problem of human head tracking is that the colors of human face and hair can vary from one person to another. When turned from the view of the camera, traditional color-based tracking systems may lose track when the initial tracking color is changed completely. In this experiment, tracking is performed under an identical environment. The color changes from the initial color of the person's face to the color of his hair. At the end, the human face will disappear completely from the camera's view. To provide for motion with such color changes, a person is directed to walk in front of the bookcases in an office while facing the camera as he starts to walk. While walking, he turns his face from right to left completely cover his face from the camera.

The motion tracking results are shown Figure 13. Sample frames in Fig. 13(a) show the tracking results using the mean shift algorithm. Sample frames in Fig. 13(b) show the tracking results using the proposed algorithm. While the experiment shows that the new algorithm provides successful tracking results, the color-based mean shift tracking algorithm lost track of the human face poses when the color changed with respect to the initial color because the algorithm is designed with the assumption of color constancy, and therefore searches for any object that satisfies the initial color condition.

Comparison Between Our Method and Automatic Updating Method

This experiment is designed to compare tracking results using two different methods: the automatic search window updating method and our method. In CDM methodology, the CDM/PDF/color histogram is updated with a damping factor in every frame.[18] The tracking object for this experiment is a white cup. To test the capability of tracking an identical

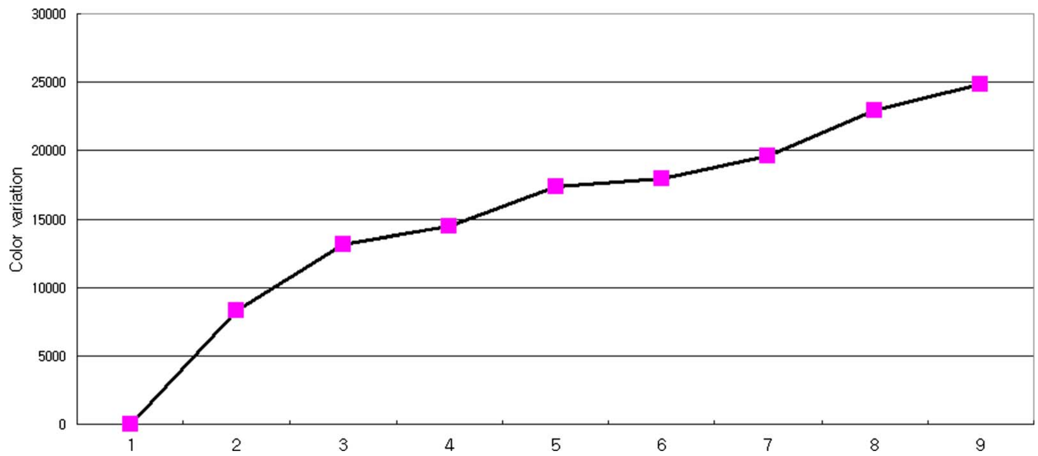

Figure 11. Color variation graph of the target object in Fig. 9.

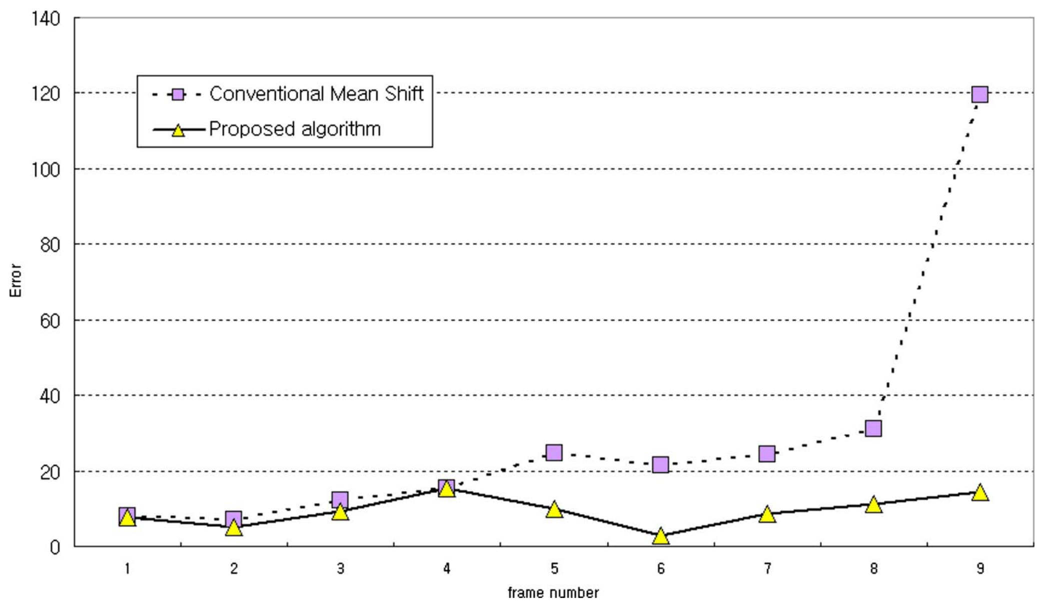

Figure 12. Plots of location tracking errors using conventional mean shift and the proposed algorithm.

target from an identical background, the moving background is designed to be as similar as possible to the target. During the experiment, the white cup is moved within close proximity to a white background. During the tracking, the automatic updating method updates the color histogram of the target from image frame to image frame without using any threshold value, while our method updates the search window using a color variation threshold value. Image frames from Figure 14(a) show the tracking results using the automatic updating method. The method can track the target over the first two frames but lost track in the third and the final image frames because of the absence of a threshold for color variation induces incorrect responses. On the other hand, the image frames in Fig. 14(b) shows the tracking results using our method. As a result of applying an appropriate color variation threshold value, this method tracks the motion of the target successfully. The color variation threshold ensures robust tracking when the target and background have a similar color distribution.

Computational Time Comparison for Two Algorithms

The comparison of the computational time using the CDM method and our method shows an advantage in tracking speed with this method.[18] The CDM method tracks a human face under varying illuminations and color-changing targets by switching color spaces using the mean shift CDM.[18] While it treats color variation of the target in a manner similar to this method, they focus on a specific target such as face tracking. Generally face tracking can easily determine color spaces such as CbCr and HS. In general object tracking, the number of color spaces in CDM method needs to be predetermined. The determination of color space to be used in the proposed algorithm is based on the sensitivity of the target color in the selected space. The entire computation time for CDM algorithm will increase proportionally to the number of the selected color spaces. The computational time of CDM and the proposed algorithm can be calculated using the following equations:

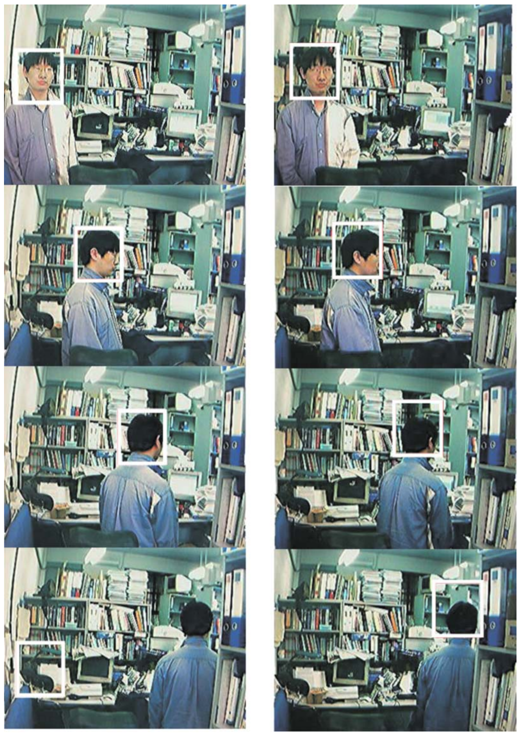

Figure 13. Experimental tracking results of human subject with radical color changes: (a) using the conventional mean shift analysis and (b) using the proposed modified mean shift algorithms.

$$\text{CDM algorithm:} \quad T_1 = NT_c + T_s + NC, \tag{3}$$

$$\text{Proposed algorithm:} \; T_2 = T_c, \tag{4}$$

where N is the number of color spaces, T_c is the mean shift computation time, T_s is the color space-switching computation time, and c is the computation time of the external search window through all frames. The proposed algorithm uses only the mean shift computational time in a single color space while the CDM algorithm computes for all color spaces. Thus, this method is more effective in terms of computation time. The algorithm can be applied to all color spaces. In our experiment, RGB color space is used to show the tracking performance because RGB color values of the object show a larger color variation than those of other color spaces.

CONCLUSIONS

In this paper, we propose a robust object tracking system which is effective for radical changes of color distribution. Conventional color-based object tracking such as mean shift requires that a target conserves a minimum amount of initial color information. However, this requirement depends on the object and its environmental elements. When an object has a large color variation, the tracking algorithm may lose the target.

We provide a robust tracking algorithm which solves this problem using a modified mean shift algorithm. The proposed algorithm detects color variation of a target object

Figure 14. Comparison of two tracking methods: (a) tracking results of the automatic updating method and (b) tracking results of this method.

by calculating difference in color distributions. Whenever the target exceeds a specific threshold, the search window based on color variation is updated iteratively. This guarantees the tracking of the target accurately even with target color variation and movement concurrently. The proposed algorithm proved its robustness with respect to these variations through experiments and comparisons. From experimental outcomes, the proposed algorithm shows successful object tracking with radical color variation in both artificial and real-world environments. We have shown that our proposed algorithm is an effective approach to tracking real objects with radical color changes. Our future work will research a robust object tracking algorithm using both color variation and shape information.

ACKNOWLEDGMENTS

This work was supported by the Korea Science and Engineering Foundation (KOSEF) grant funded by the Korea government (MOST) (No. R01-2005-000-10940-0).

REFERENCES

[1] I. Matthews, T. Ishikawa, and S. Baker, “The template update problem”, IEEE Trans. Pattern Anal. Mach. Intell. **26**(6), 810–815 (2004).
[2] N. Gordon and D. Salmond, “Bayesian state estimation for tracking and guidance using the bootstrap filter”, J. Guid. Control Dyn. **18**(6), 1434–1443 (1995).
[3] K. Nummiaro, E. Koller-Meier, and L. V. Gool, “An adaptive color-based particle filter, J. Image and Vision Computing”, Lect. Notes, Comput. Sci. **21**(1), 99–110 (2003).
[4] A. Jepson, D. Fleet, and T. El-Maraghi, “Robust online appearance

models for visual tracking", *Proc. IEEE Conf. on Computer Vision and Pattern Recognition* (IEEE, Los Alamitos, CA, 2001) pp. 415–422.

[5] S. McKenna, Y. Raja, and S. Gong, "Tracking color objects using adaptive mixture models", J. Image Vis Comput. **17**(3), 223–229 (1999).

[6] K. Fukanaga and L. Hostetler, "The estimation of the gradient of a density function, with applications in pattern recognition", IEEE Trans. Inf. Theory **21**(1), 32–40 (1975).

[7] Y. Cheng, "Mean shift, mode seeking, and clustering", IEEE Trans. Pattern Anal. Mach. Intell. **17**(8), 790–799 (1995).

[8] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis", IEEE Trans. Pattern Anal. Mach. Intell. **24**(5), 603–619 (2002).

[9] D. Comaniciu, V. Ramesh, and P. Meer, "Kernel-based object tracking", IEEE Trans. Pattern Anal. Mach. Intell. **25**(5), 564–575 (2003).

[10] D. Comaniciu and V. Ramesh, "Mean shift and optimal prediction for efficient object tracking", *Proc. IEEE Int. Conf. on Image Processing* (IEEE, Los Alamitos, CA, 2000) pp. 70–73.

[11] R. Collins, "Mean-shift blob tracking through scale space", *Proc. IEEE Conf. on Computer Vision and Pattern Recognition* (IEEE, Los Alamitos, CA, 2003) pp. 234–240.

[12] D. Comaniciu, V. Ramesh, and P. Meer, "Real-time tracking of non-rigid objects using mean shift", *Proc. IEEE Conf. on Computer Vision and Pattern Recognition* (IEEE, Los Alamitos, CA, 2000), pp. 142–149.

[13] F. Porikli and O. Tuzel, "Human body tracking by adaptive background models and mean shift analysis", *IEEE Int. Workshop on Performance Evaluation of Tracking and Surveillance* (IEEE, Piscataway, NJ, 2003).

[14] R. V. Babu, P. Perez, and P. Bouthemy, "Robust tracking with motion estimation and kernel-based color modeling", *Proc. IEEE Int. Conf. on Image Processing* (IEEE, Los Alamitos CA, 2005) pp. 717–720.

[15] D. Toth, I. Stuke, A. Wagner, and T. Aach, "Detection of moving shadows using mean shift clustering and a significance test", *Proc. IEEE Conf. on Pattern Recognition* (IEEE, Los Alamitos, CA, 2004) pp. 260–263.

[16] D. Comaniciu and P. Meer, "Robust analysis of feature spaces: Color image segmentation", *Proc. IEEE Conf. on Computer Vision and Pattern Recognition* (IEEE, Los Alamitos, CA, 1997) pp. 750–755.

[17] Y. Cheng, "Mean-shift, mode seeking, and clustering", IEEE Trans. Pattern Anal. Mach. Intell. **17**(8), 790–799 (1995).

[18] H. Stern and B. Efros, "Adaptive color space switching for tracking under varying illumination", J. Image Vis Comput. **23**(3), 353–364 (2005).