Controlling a Camera with Minimized Camera Motion Changes under the Constraint of a Planned Camera-work

Yasutaka ATARASHI†, Yoshinari KAMEDA††, Masayuki MUKUNOKI††, Koh KAKUSHO††,
Michihiko MINOH†† and Katsuo IKEDA‡

† Graduate School of Informatics, Kyoto University
†† Center for Information and Multimedia Studies, Kyoto University
‡Osaka Institute of Technology

Abstract

We propose a method to control the camera in order to obtain video images that have minimum changes of camera motion under the constraint of a planned camera-work. A camera-work for shooting an object can be defined by the position and the size of the object in each video image. In order to obtain video images that are comfortable for a human to see, it is not necessary that the position and the size of the object in the images are exactly the same as the specified camera-work, but it is important that the changes of the camera motion is small. Considering this issue, we control the camera to minimize the changes of the camera motion so that the position and the size of the object continue to be within the acceptable range.

1. Introduction

Various studies have been presented in order to take videos of human activities such as lectures, music concerts, sports games and so on automatically[1][2]. It is required to shoot a moving object for taking those videos.

We consider that the shooting is specified by a value of a camera-work. A camera-work is defined as a vector that consists of the position, the velocity, the size and the magnification rate of the object, and the velocity of the background in a video image at each frame. We call the specified value of the camera-work a “target camera-work” (TCW). All the elements in the TCW do not need to be specified. For example, when we want to shoot a lecturer at the center of the image in medium size, only the position and the size of the object are specified.

It is required to control the camera by referring images since it sometimes happens that the object goes out of the image due to uncontrollable factors such as noise in measuring the location of the object and unexpected object motions even if the camera is controlled based on the current 3D position of the object in the environment.

Visual servo [3] is conventional technique to control the camera by referring images. The technique always controls the camera to reduce the difference of the current resultant camera-work in the image from the TCW. When we employ this technique, the camera motion is not stabilized but continuously modified to realize the TCW exactly.

In the case we take videos of human activities so that the videos are comfortable for human to see, it is important to reduce not only the difference of the resultant camera-work from the TCW but also the changes of the camera motion. The changes of the camera motion cause the unstable motion of the background in the taken images. It is not always required that the resultant camera-work in the images are exactly the same as the TCW.

In our approach, we introduce an “acceptable range” in order to minimize the changes of the camera motion. It is sufficient if the resultant camera-work is within a certain range of the TCW. The range is given as the upper bound and the lower bound for each specified element in the TCW. We consider the resultant camera-work is within the range even if the element not specified in the TCW takes any value in the resultant camera-work. This range is called the “acceptable range”.

Our process constitutes iterative steps. For each step, we estimate the resultant camera-work at the next step, which is called “estimated camera-work” (ECW). If the ECW is within the acceptable range, we do not change the camera control parameters. If not, we modify the parameters. The parameters are determined from the ECWs estimated for the next several steps, so that the ECW is within the acceptable range at as many steps in a row as possible after the modification.

In order to realize the process, the system first extracts the object region from the images by applying an M-estimator to the optical flow, and evaluates whether the

ECW is within the acceptable range by applying Kalman filter.

In the remainder of this paper, we will describe this process in detail. In section 2, we will describe the overview of the process in the proposed system. In section 3, we will describe the detail algorithm of the process. In section 4, we will present the result of a preliminary experiment. In section 5, we will give our concluding remarks.

2. Overview of the process

2.1. Environment

We consider the case that a camera shoots a single moving object. The shooting camera is modeled as a pin-hole camera with the center of projection at the center of rotation. The camera is controlled by specifying its pan/tilt speed C^p, C^t and the zoom parameter C^z. These parameters take discrete values. The focal length of the camera f is given by function $f = g(C^z)$. Since we are interested especially in shooting an indoor human, for example, a lecturer in a lecture room, it is assumed that the object is in an unknown rigid motion without rotation around the optical axis of the camera.

The camera-work is defined by the position(x,y), the velocity($\dot{x}$,$\dot{y}$), the size(s) and the magnification rate($\dot{s}$) of the object and the velocity($\dot{d}$,$\dot{e}$) of the background in the video image at each frame(Fig.1). The position of the object is the centroid of the object region. The size of the object is the width of the object, because our target human in an image is standing and moving.

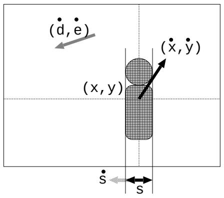


Figure 1. Elements of the camera-work

2.2. General framework of the process

General framework of the process realized by our system is shown in Fig.2.

It is assumed that the TCW is preliminary given or is planned by other systems or human. All the elements in the TCW do not need to be specified. The acceptable range is also assumed to be given in advance.

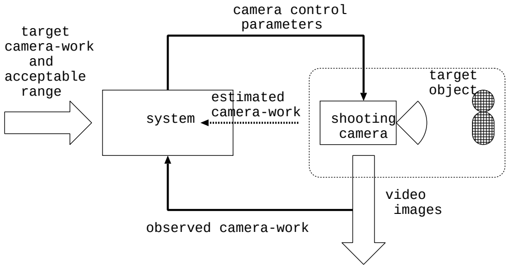


Figure 2. Framework

An observation stage, an estimation stage and a control stage constitute a step of the iterative process of the system. At the observation stage, we first extract the resultant camera-work from the current image. We call it the “observed camera-work” (OCW). At the estimation stage, we estimate the ECWs for next several steps based on the current and the previous OCWs.

At the control stage, if necessary, we modify the camera control parameters based on the ECWs so that the changes of the camera motion will be minimized.

2.3. Minimizing the changes of the camera motion

In order to minimize the changes of the camera motion, we should not modify the camera control parameters as long as possible once they are changed. The camera control parameters are changed only when the ECW at the next step is out of the acceptable range.

We determine the new camera control parameters so that the ECW is within the acceptable range at as many steps in a row as possible. We estimate the ECWs not only at next step but also for next several steps by Kalman filter. As a result, we can determine the parameters that lead to the longest period until the ECW will be out of the acceptable range again. If we would control the camera by referring the ECW only at next step, the resultant camera-work could be out of the acceptable range again soon.

The flow of the process described above is summarized below:

1. Extract the OCW from video images (as described section 3.1).

2. Improve the current ECW by updating Kalman filter using the OCW.

3. Estimate the ECW at next step based on the current ECW by applying Kalman filter.

4. If the ECW is within the acceptable range, the camera control parameters keep current values. Go to 1.

5. Otherwise, estimate the ECWs for next N_{max} steps by applying Kalman filter. The constant N_{max} is decided from the precision and the processing time of the estimation.

6. Determine the values of the camera control parameters so that the modified values can keep the ECW within the acceptable range at as many steps in a row as possible.

7. Modify the camera motion with changing the camera control parameters to the determined values.

8. Go to 1.

3. Implementation

To realize the process described in the previous section, it is required to extract the OCW and to estimate the ECWs. This section describes those implementation and how to determine the camera control parameters.

3.1. Extracting the OCW

It is required to extract the object region in order to extract the OCW. A method of background subtraction is not suitable for extracting the object region because the background would be change as the result of the pan, tilt and zoom of the shooting camera.

We extract the object region based on the optical flow. The optical flow is different between the object region and the background region when the object is moving. We can extract the object region by classifying the optical flow into 2 classes, i.e. the object and the background, and selecting one of them.

The classification is done by applying M-estimator to the optical flow (it will be described in section 3.1.1). When the object is not moving, we cannot extract the object region. In such case, the position of the object is not calculated and we use the position in the current ECW as the current position of the object.

The selection of the region is done by comparing the distance between the extracted motion parameters of the region and the expected motion parameters of the background region. The expected parameters are calculated from previous values of the camera control parameters (it will be described in section 3.1.2).

The flow of the process to extract the OCW is shown below:

1. Calculate optical flow(u_i^p,v_i^p) at the point that has coordinate (x_i^p,y_i^p) in an image for $i = 1, \cdots, N_{opt}$ where N_{opt} is the number of points that optical flow are calculated.

2. Apply the M-estimator to the optical flow to estimate motion parameters of the first region. The motion parameters represent the motion of the region. The motion consists of a horizontal component(H), a vertical component(V) and a scale component(S). The motion parameters of the first region are denoted by (H_F,V_F,S_F).

3. Classify (x_i^p,y_i^p) as the first region if $|H_F + S_F x_i - u_i| < \Theta_{seg}$ and $|V_F + S_F y_i - v_i| < \Theta_{seg}$ where Θ_{seg} is the threshold of the classification.

4. Compare $N_{opt} - N_F$ with Θ_{reg} where N_F denotes the number of the points classified as the first region and Θ_{reg} is the threshold.

   (a) If $N_{opt} - N_F < \Theta_{reg}$, it is considered that the object is not moving. We extract the velocity of the object as the optical flow at the previous estimated position of the object, and the velocity of the background as the optical flow at the image center. The process of the extraction ends here.

   (b) If not, apply the M-estimator to the optical flow at points that are not classified as the first region in order to estimate motion parameters of the other region denoted by (H_O, V_O, S_O).

5. Calculate the expected parameters (H_r,V_r,S_r) of the background region from the previous camera control parameters.

6. If (H_F, V_F, S_F) is further from (H_r, V_r, S_r) than (H_O, V_O, S_O), decide that the first region is the object region. If not, the other region is the object region.

7. Extract all the elements in the OCW. From the extracted object region, we extract the position of the object as the centroid, the velocity of the object as the optical flow at the centroid and the size of the object as the width. The magnification rate of the object is extracted as the extracted size of the object times $S_* - 1$ (S_* represents either S_F or S_O) and the velocity of the background as the optical flow at the image center.

3.1.1. Estimating motion parameters

The motion parameters are estimated in order to classify the optical flow.

There is no rotation around optical axis and no camera translation. Now optical flow(u_i^p,v_i^p) at the point (x_i^p,y_i^p) is given by:

$$u_i^p = H + S x_i^p, \quad v_i^p = V + S y_i^p \qquad (1)$$

We introduce M-estimator to estimate the motion parameters (H, V, S). M-estimator is one of popular robust statistics techniques. We can estimate the motion parameters by M-estimator without the influence of outliers. The result of calculating the optical flow often includes outliers. When we estimate the motion parameters of the first region at the step 2 in extracting the OCW, the optical flow calculated at the point in the other region can also act as the outliers.

Introducing M-estimator, we minimize the function defined below to obtain the motion parameters (H,V,S).

$$\sum_i \rho(u_i^p - H - Sx_i^p) + \sum_i \rho(v_i^p - V - Sy_i^p) \tag{2}$$

where

$$\rho(z) = \log\left(1 + \frac{1}{2}\left(\frac{z}{\sigma}\right)^2\right) \tag{3}$$

We use the Lorentzian function as the weight function ρ and the parameter σ is given by $\sigma = \frac{\Theta_{seg}}{\sqrt{2}}$[4].

3.1.2. Selection of the motion parameters of the object

If two regions are extracted, we need to select either region as the object region. The region whose values of the motion parameters are further from the expected values of the background region is selected.

The expected values (H_r, V_r, S_r) are calculated from the previous values of the camera control parameters. They are given by:

$$H_r = -\Delta g(C_{n-1}^z)C_{n-1}^p \tag{4}$$

$$V_r = -\Delta g(C_{n-1}^z)C_{n-1}^t \tag{5}$$

$$S_r = \frac{g(C_{n-1}^z) - g(C_{n-2}^z)}{g(C_{n-2}^z)} \tag{6}$$

where the values at step n is denoted by subscript “n”, and Δ is the time interval between the steps, and $g(C_n^z)$ is the focal length corresponding C_n^z.

We define a distance between motion parameters (H_r, V_r, S_r) and (H, V, S) as:

$$\begin{aligned} &\int_x |(H_r + S_r x) - (H + Sx)|dx \\ &+ \int_y |(V_r + S_r y) - (V + Sy)|dy \end{aligned} \tag{7}$$

We select the region with longer distance from (H_r,V_r,S_r) as the object region.

We use the calculated motion parameters not for eliminating the optical flow caused by the camera motion but for selecting the object region. This is because the calculated motion parameters are not always accurate due to the characteristics of the mechanical part of the camera.

3.2. Estimating the ECWs by Kalman filter

We estimate the ECWs by Kalman filter.

It is assumed that the motion of the object is in uniform accelerated motion in the image and the camera motion has uniform angular velocity until the camera control parameters are changed. These are reasonable assumptions compared with information we can obtain from images, i.e. the OCW. We cannot effectively use more detailed models due to the limitation of the information.

To derive Kalman filter, we define the system equation and the observation equation.

The system equation is defined as:

$$\boldsymbol{\alpha}_{n+1} = \boldsymbol{A\alpha} + \boldsymbol{c}_n + \boldsymbol{\eta}_n \tag{8}$$

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{U} & & & & \\ & \boldsymbol{U} & & & \\ & & \boldsymbol{U} & & \\ & & & 1 & \\ & & & & 1 \end{bmatrix} \tag{9}$$

$$\boldsymbol{U} = \begin{bmatrix} 1 & \Delta & \frac{1}{2}\Delta^2 \\ 0 & 1 & \Delta \\ 0 & 0 & 1 \end{bmatrix} \tag{10}$$

The state vector $\boldsymbol{\alpha}_n$ is given by:

$$\boldsymbol{\alpha}_n = \begin{bmatrix} x_n^{rst} & \dot{x}_n^{rst} & \ddot{x}_n^{rst} & y_n^{rst} & \dot{y}_n^{rst} & \ddot{y}_n^{rst} & s_n^{rst} & \dot{s}_n^{rst} & \ddot{s}_n^{rst} & \dot{d}_n^{rst} & \dot{e}_n^{rst} \end{bmatrix}' \tag{11}$$

where x_n^{rst}, $\dot{x}_n^{rst}$, y_n^{rst}, $\dot{y}_n^{rst}$, s_n^{rst}, $\dot{s}_n^{rst}$, $\dot{d}_n^{rst}$ and $\dot{e}_n^{rst}$ are the resultant camera-work at step n, and $\ddot{x}_n^{rst}$, $\ddot{y}_n^{rst}$ and $\ddot{s}_n^{rst}$ are accelerations of x_n^{rst},y_n^{rst} and s_n^{rst}, respectively. The control vector $\boldsymbol{c}_n$ is given by:

$$\boldsymbol{c}_n = \begin{bmatrix} -\Delta g(C_{n-1}^z)\delta C^p + \lambda x_n^{rst} \\ -g(C_{n-1}^z)\delta C^p + \lambda \dot{x}_n^{rst} \\ \lambda \ddot{x}_n^{rst} \\ -\Delta g(C_{n-1}^z)\delta C^t + \lambda y_n^{rst} \\ -g(C_{n-1}^z)\delta C^t + \lambda \dot{y}_n^{rst} \\ \lambda \ddot{y}_n^{rst} \\ \lambda s_n^{rst} \\ \lambda \dot{s}_n^{rst} \\ \lambda \ddot{s}_n^{rst} \\ -g(C_{n-1}^z)\delta C^p + \lambda \dot{d}_n^{rst} \\ -g(C_{n-1}^z)\delta C^t + \lambda \dot{e}_n^{rst} \end{bmatrix} \tag{12}$$

where

$$\lambda = \frac{g(C_n^z) - g(C_{n-1}^z)}{g(C_{n-1}^z)} \tag{13}$$

$$\delta C^p = C_n^p - C_{n-1}^p \tag{14}$$

$$\delta C^t = C_n^t - C_{n-1}^t \tag{15}$$

The system noise vector $\boldsymbol{\eta}_n$ has average **0** and a given covariance matrix $\boldsymbol{Q}$.

The observation equation is defined as:

$$\boldsymbol{\beta}_n = \boldsymbol{B}_n \boldsymbol{\alpha}_n + \boldsymbol{\epsilon}_n \tag{16}$$

where $\boldsymbol{\beta}_n$ is the observation vector, that is the OCW at step n, given by:

$$\boldsymbol{\beta}_n = \left[x_n^{obs} \ \dot{x}_n^{obs} \ y_n^{obs} \ \dot{y}_n^{obs} \ s_n^{obs} \ \dot{s}_n^{obs} \ \dot{d}_n^{obs} \ \dot{e}_n^{obs}\right] \quad (17)$$

The matrix $\boldsymbol{B}_n$ is given by:

$$\boldsymbol{B}_n = \begin{bmatrix} a_n & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_n & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & a_n & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (18)$$

where a_n is 1 if we can extract the object region, and 0 if not.

The observation noise vector $\boldsymbol{\epsilon}_n$ has average $\mathbf{0}$ and a given covariance matrix $\boldsymbol{H}$.

Kalman filter is implied by the equations from (8) to (18) as follows:

$$\begin{aligned} \boldsymbol{\alpha}_{n+k|n} &= \boldsymbol{A}\boldsymbol{\alpha}_{n+k-1|n} + \boldsymbol{c}_{n+k-1} && (19) \\ \boldsymbol{\Sigma}_{n+1|n} &= \boldsymbol{A}\boldsymbol{\Sigma}_{n|n}\boldsymbol{A}' + \boldsymbol{Q} && (20) \\ \boldsymbol{G}_n &= \boldsymbol{\Sigma}_{n|n-1}\boldsymbol{B}_n'(\boldsymbol{B}_n\boldsymbol{\Sigma}_{n|n-1}\boldsymbol{B}_n' + \boldsymbol{H})^{-1} && (21) \\ \boldsymbol{\alpha}_{n|n} &= \boldsymbol{\alpha}_{n|n-1} + \boldsymbol{G}_n\{\boldsymbol{\beta}_n - \boldsymbol{B}_n\boldsymbol{\alpha}_{n|n-1}\} && (22) \\ \boldsymbol{\Sigma}_{n|n} &= (\boldsymbol{I} - \boldsymbol{G}_n\boldsymbol{B}_n)\boldsymbol{\Sigma}_{n|n-1} && (23) \end{aligned}$$

where $\boldsymbol{\alpha}_{i|j}$ is estimated $\boldsymbol{\alpha}_i$ based on observations from step 1 to step j, $\boldsymbol{\Sigma}_{i|j}$ is estimated covariance matrix of the error at step i based on observation from step 1 to step j, and $\boldsymbol{G}_n$ is the Kalman gain matrix, and the transpose of the matrix is denoted by superscript "$'$".

After the observation at step n, the OCW gives $\boldsymbol{\beta}_n$. Matrices $\boldsymbol{\Sigma}_{n-1|n-1}$, $\boldsymbol{\Sigma}_{n|n-1}$ and $\boldsymbol{G}_n$ are calculated from the equations (23), (20) and (21) in this order. Then, the current ECW in $\boldsymbol{\alpha}_{n|n}$ is calculated from the equation (22). As the result, we can calculate the ECWs for next N_{max} steps that are included in $\boldsymbol{\alpha}_{n+k|n}$ ($k = 1, 2, \cdots, N_{max}$) from the equation (19).

3.3. Controlling camera

We modify the camera control parameters if the ECW at next step is out of the acceptable range. When the modification is required, we calculate the values of the camera control parameters so that the ECWs is within the acceptable range at as many steps in a row as possible.

The camera control parameters C_n^z, C_n^p and C_n^t are determined from the ECWs. We first determine C_n^z from s and $\dot{s}$, then C_n^p from x, $\dot{x}$ and $\dot{d}$, and C_n^t from y, $\dot{y}$ and $\dot{e}$. The position and the velocity of the object are influenced by all the camera control parameters. The size and the magnification rate of the object are influenced by C_n^z only. Therefore, C_n^z is uniquely determined from the size and the magnification rate of the object.

We determine C_n^z as described below.

If s and $\dot{s}$ are not specified in the TCW, since the ECW is within the range even if s and $\dot{s}$ in the ECW take any values, we do not modify C_n^z. If s and/or $\dot{s}$ is specified in the TCW, we calculate the range of C_n^z that s and $\dot{s}$ in the ECW at step $(n+k)$ are between the upper bound and the lower bound specified by the acceptable range for $k = 1, 2, 3, \cdots, N_{max}$. The range of C_n^z is calculated from the relation between the ECW at step $(n+k)$ and C_n^z. In order to obtain the relation, we obtain the relation between $\boldsymbol{\alpha}_{n+k|n}$ and $\boldsymbol{c}_n$. This is because the ECW at step $(n+k)$ is included in $\boldsymbol{\alpha}_{n+k|n}$ and C_n^z is referred in $\boldsymbol{c}_n$. The relation is given from the equation (19) by:

$$\boldsymbol{\alpha}_{n+k|n} = \boldsymbol{A}^k\boldsymbol{\alpha}_{n|n} + \boldsymbol{A}^{k-1}\boldsymbol{c}_n \quad (24)$$

We can calculate $\boldsymbol{\alpha}_{n|n}$ after the observation at step n.

Therefore, we can calculate $\boldsymbol{c}_n$ that will cause the camera-work in $\hat{\boldsymbol{\alpha}}_{n+k}$ at step $(n+k)$ by:

$$\boldsymbol{c}_n = \boldsymbol{A}^{-(k-1)}\left(\hat{\boldsymbol{\alpha}}_{n+k} - \boldsymbol{A}^{-k}\alpha_{n|n}\right) \quad (25)$$

We calculate C_n^z from $\boldsymbol{c}_n$ by the equation (12).

Then, we take the intersection of the range of C_n^z from step $(n+1)$ to step $(n+N_{max})$. We determine the value of C_n^z as the average between the upper bound and the lower bound of the intersection.

After C_n^z is determined, we determine C_n^p and C_n^t in the similar way as C_n^z is determined. We control the camera with the determined values of C_n^p, C_n^t and C_n^z.

4. Experiment

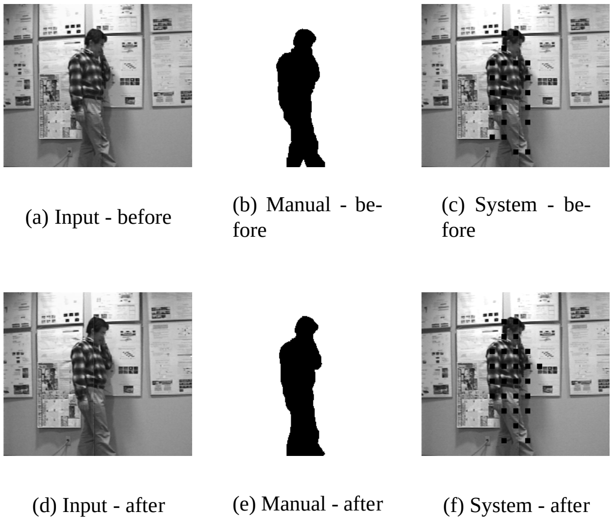
We conducted a preliminary experiment to evaluate the extraction of the OCW and the estimation of the ECWs by Kalman filter when shooting a real moving object. The TCW is given so that the object is at the horizontal center in the images. The acceptable range is not yet introduced.

The size of the input image is 256×220 pixels. Optical flow is calculated at 150 points(15 columns × 10 rows) by block matching. The object is a human walking with approximately constant speed along a straight line at 3 meters from the camera.

The OCW is extracted by the method described in section 3.1. Examples of input image are shown in Fig.3(a) and (d). The object regions extracted manually for Fig.3(a) and (d) are shown in Fig.3(b) and (e), respectively. The object regions extracted by the method for Fig.3(a) and (d) are shown in Fig.3(c) and (f), respectively. A black rectangle shows that the point is classified as the object region. The extracted OCW between Fig.3(a) and (d) is shown in Table.1. The unit in the table is a pixel.

Table 1. Observed camera-work

	x^{obs}	$\dot{x}^{obs}$	y^{obs}	$\dot{y}^{obs}$	s^{obs}	$\dot{s}^{obs}$	$\dot{d}^{obs}$	$\dot{e}^{obs}$
manual	119.6	2.2	128.9	-0.8	61	-4	-2.3	0.3
system	122.3	3.3	114.6	0.4	64	0.6	-1.3	0.1



(a) Input - before

(b) Manual - before

(c) System - before

(d) Input - after

(e) Manual - after

(f) System - after

Figure 3. Extraction results

The optical flow is calculated at horizontal intervals of about 17 pixels and vertical intervals of 22 pixels. The resolution of extracting the position and the size of the object depends on the intervals. The resolution of extracting the velocity of the object and the background depends on the unit of calculating the optical flow. The result shows that we can extract the OCW with reasonable precision compared with the resolution.

The extracted position of the object is more inaccurate than the extracted velocity. It is because the number of the points that optical flow is calculated is small. If we increase the number, the results will be more accurate. At the same time, it takes more time to calculate the optical flow and as the result the camera control is delayed. A trade-off problem between the resolution and the processing time exists. The extracted size of the object is also inaccurate because of the error of the classification between the object region and the background region. We need to improve the classification method to be more robust by, for example, considering the result of the classification at the neighborhood or at the previous step.

Fig.4 shows x in the OCW and in the ECWs at every 10 steps for the next 10 steps. The time interval(Δ) between steps is 100 millisecond. Though the OCW does not always show the resultant camera-work due to the error of the observation, it gives an indication of the resultant camera-work.

The ECWs estimated at step 40 are far from the OCW between step 42 and 50. This is because the camera motion is often changed in this experiment. If we introduce the acceptable range, the ECW is expected to be closer to the OCW.

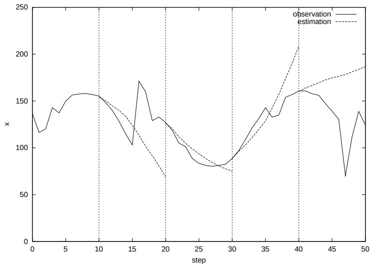


Figure 4. Estimated camera-work

5. Conclusion

We proposed the method to control the camera in order to take videos with the minimum changes of the camera motion under the constraint of the TCW.

In our approach, we introduce the “acceptable range” in order to minimize the changes of the camera motion. For each step, we estimate the resultant camera-work at the next step. If the ECW is within the acceptable range, we do not change the camera control parameters. If not, we modify the parameters. The parameters are determined from the ECWs estimated for the next several steps, so that the ECW is within the acceptable range at as many steps in a row as possible after the modification.

We implemented the method to extract the OCW and to estimate the ECWs, and conducted the preliminary experiment. The result of the experiment shows that the OCW is extracted and the ECWs are estimated with reasonable precision when we shoot a real moving object.

We will implement the whole proposed method, and conduct the experiment to evaluate the effectiveness of the method.

References

[1] Q.Cai and J.K.Aggarwal. Real time tracking for enhanced tennis broadcasts. In *Proceedings Computer Vision and Pattern Recognition (CVPR’98)*, pages 68–72, 1998.

[2] A. G. Q.Liu, Young Rui and J. Cadiz. Automating camera management for lecture room environments. In *ACM CHI 2001*, 2001.

[3] S.A.Hutchinson, G.D.Hager, and P.I.Corke. A tutorial on visual servo control. *IEEE Trans. Robot. Automat.*, 12(5):651–670, 1996.

[4] M.J.Black. The robust estimation of multiple motions. *Computer Vision and Image Understanding*, 63(1):75–104, 1996.