Kalman Filter-Based Coarse-to-Fine Control for Display Visual Alignment Systems

SangJoo Kwon*, *Member, IEEE*, Haemin Jeong, and Jaewoong Hwang

***Abstract*—A coarse-to-fine two-stage control method is investigated for the display visual alignment systems. The proposed visual servo is with hierarchical loops, where the original fine but slow vision loop is necessary for the exact localization of alignment marks while the coarse but fast vision loop of exploiting pruned image data is to compensate for the mask-panel misalignment. The degraded resolution of the reduced images is recovered in terms of the Kalman filter which tracks the mark centroids in near real-time. In order to construct the recursive estimation algorithm, the motion model for the moving alignment marks is determined by solving the forward kinematics of positioning mechanism and the measurements from vision sensors are given by means of the geometric template matching (Kwon and Hwang, "Kinematics, pattern recognition, and motion control of mask-panel alignment system," *Control Eng. Practice*, vol. 19, pp. 883–892, 2011). Compared with the conventional alignment methods, this approach enables a fast and fine alignment control. Experimental results are followed to validate the proposed control framework.**

***Note to Practitioners*—In order to successfully apply the developed alignment control to any display manufacturing equipment, it is necessary to well understand the principle of the geometric template matching (GTM) as an alignment mark specific fast algorithm, the details on which can be consulted in our preceding works (Kwon and Hwang, "Kinematics, pattern recognition, and motion control of mask-panel alignment system," *Control Eng. Practice*, vol. 19, pp. 883–892, 2011). The new approach has the goal of updating the pose of an alignment mark as fast as the capturing rate of a frame grabber by utilizing pruned image data but recovering the lost resolution in terms of the Kalman filter. For example, in using a common 30 fps grabber, the reduced image of 320 $\times$ 240 pixels is a proper choice to finish the image processing and Kalman filtering within 30 ms under GTM. The proposed algorithm can be implemented in the current industrial display aligners by modifying the control software so that the reference inputs for the distributed joint servos follow the error compensation trajectory in Fig. 4 with the Kalman filter estimates.**

***Index Terms*—Flat panel display, image processing, Kalman filter, observer-based control, visual alignment.**

I. Introduction

In THE flat panel display industry, the mask-panel visual alignment system is operating to correct the misalignment between mask and panel prior to performing a specific manufacturing process. The efficiency of the autonomous visual alignment system is a critical problem to determine the productivity in any manufacturing equipment which includes the mask-panel alignment process as well as in the display industry [1]–[5]. Although the visual alignment systems are widely applied throughout the display and semiconductor production lines, the alignment control technology is open to further development to improve the performance and the reliability.

In general, the visual alignment system consists of the vision system to recognize the alignment marks on the mask and panel and the motion control system to compensate the misalignment between the mask and panel by moving multi-axis positioning stage. In the vision system, at least two CCD cameras are involved to capture the alignment mark images and sequential image processing is followed for the raw image data to find the mark positions. And then the quantity of mask-panel misalignment is determined and the stage control system is activated to compensate the alignment error according to appropriate reference trajectory and feedback control logic. The alignment process will be finished when the final misalignment meets a given tolerance.

Hence, the overall performance of a visual alignment system depends not only on the efficiency of the image processing algorithm related to how to look at the alignment marks through the vision system but also on the motion control strategy concerned with how to move the positioning mechanism to compensate the misalignments. In improving the visual alignment performance, the first consideration would be to adopt high-speed cameras with high pixel resolution CCD sensors and also a high precision positioning mechanism. But, the camera speed is meaningless as far as the image processing time is over the capturing period and the high resolution images would require as much increase of the image processing time. Moreover, these are directly connected to the cost increase. Here, in order to make a progress in the visual alignment technology, it is earnestly desired to seek an efficient way to best utilize given sensory information in both image processing and motion control.

In fact, the visual alignment performance is greatly dependent upon how fast the misalignment states can be updated by means of the exteroceptive vision sensors. However, since the capturing rate of the frame grabber is limited and an amount of time is necessary for object recognition, there exists a certain limit in increasing the frequency of visual feedback. To address this problem, prior to this paper, we have suggested a geometric template matching (GTM) in [3] and [4] as an alignment mark specific fast vision algorithm, where it was shown that it enables an order of reduction in the mark recognition time and it can be applied as a practical alternative to time-consuming vision libraries. As a follow on work, this paper deals with an advanced but practical motion control method for the visual alignment systems.

The observer-based control has been a hot issue for a long time, which is purposed to overcome the lack or the uncertainty in sensory information by using a filtering framework for sensor fusion. For example, the recent works on tire friction estimation and braking control [6], wind angle estimation for autopilot design [7], and decentralized output feedback control for industrial boilers [8] demonstrate the benefit of observer-based control in the specific applications. Also, the coarse-to-fine strategy is very popular to get over the bandwidth limit in motion control and image processing [9]–[15].

This paper proposes an observer-based coarse-to-fine control for the display visual alignment systems. Due to the long image processing time, most industrial alignment systems are based on the look-then-move control where stage control and image processing are sequentially conducted. Then, it is greatly disadvantageous in precision control because it would give rise to microscopic stick-slip motions. To the contrary, the suggested approach achieves a near real-time visual feedback in terms of the fast image processing of GTM and the two-stage vision loops, which enables a great reduction in the alignment time. In addition, the Kalman filter [16] plays a crucial role as a state observer

Manuscript received November 30, 2010; revised November 21, 2011; accepted April 18, 2012. This paper was recommended for publication by Associate Editor C.-F. Chien and Editor K. Goldberg upon evaluation of the reviewers' comments. This work was supported by the Korea National Research Foundation under contract number 2009-0077203. This was presented in part at the 2009 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM), Singapore, July 14–17, 2009. *Asterisk indicates corresponding author.*

*S. J. Kwon is with the School of Aerospace and Mechanical Engineering, Korea Aerospace University, Goyang 412-791, Korea (e-mail: sjkwon@kau.ac.kr).

H. Jeong and J. Hwang are with the School of Aerospace and Mechanical Engineering, Korea Aerospace University, Goyang 412-791, Korea (e-mail: {seamanjhm,bluelead}@kau.ac.kr, +82-2-300-0178).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TASE.2012.2196693

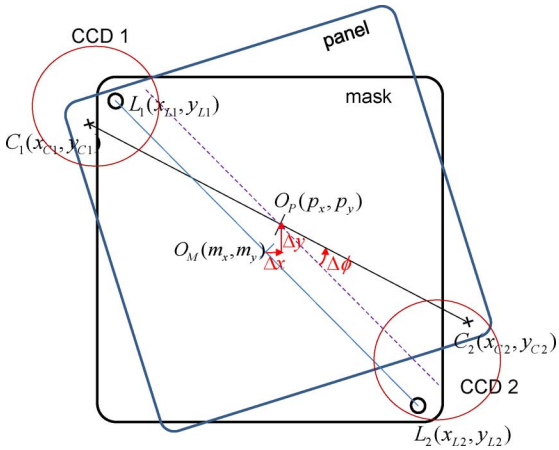

Fig. 1. Alignment marks and the in-plane mask-panel misalignment.

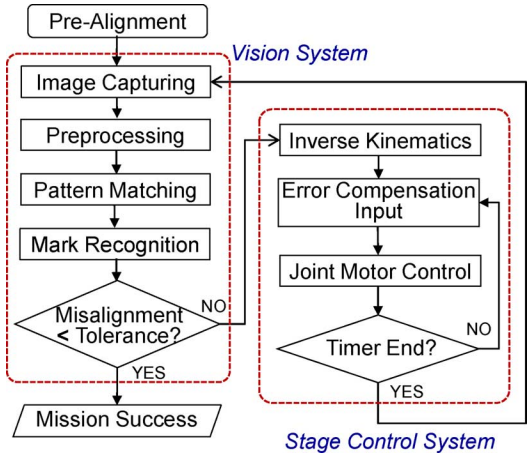

Fig. 2. Overall sequence of the visual alignment process.

to recover the lost resolution of the reduced images. This paper is a revised and extended version of the prior work [5], where a lot of parts have been technically refined with new experiment results.

II. Visual Alignment Process in Display Industry

When a glass panel is transferred into the station (or chamber) of a display manufacturing equipment, the visual alignment system is activated so that the pattern mask is aligned to the glass panel prior to the main process such as evaporation, printing, and lithography. As the schematic shows in Fig. 1, normally two CCD cameras are used to localize the alignment marks on the mask and panel. Once the centroids of marks (L_1, L_2) of the mask and (C_1, C_2) of the panel are given as a result of vision processing, it is straightforward to determine the in-plane misalignment between mask and panel as follows [3].

$$\begin{aligned}\Delta x &= p_x - m_x = \frac{(x_{C1} + x_{C2} - x_{L1} - x_{L2})}{2} \\ \Delta y &= p_y - m_y = \frac{(y_{C1} + y_{C2} - y_{L1} - y_{L2})}{2} \\ \Delta\phi &= \tan^{-1}\left(\frac{y_{C2} - y_{C1}}{x_{C2} - x_{C1}}\right) - \tan^{-1}\left(\frac{y_{L2} - y_{L1}}{x_{L2} - x_{L1}}\right). \end{aligned} \quad (1)$$

The image processing time to get the alignment mark poses is varied depending on the image size of the vision sensor and the microprocessor used. But, it is known approximately hundreds of milliseconds when the general matching algorithm is applied. Due to the large image processing time, the visual alignment system is usually operated in a look-then-move manner as the overall flow is denoted in Fig. 2, where the prealignment process makes the alignment marks on the mask and panel enter into the field of views of the vision cameras. Then, the images are captured by the frame grabber and the centroidal postures of the marks are found through the pattern matching algorithm.

When the mask-panel misalignment in (1) is given in the task space, it can be compensated for by the controlled motion of the alignment stage which mounts the pattern mask on the moving platform. The compensation displacements of joint actuators can be determined in terms of the inverse kinematics solution for the adopted alignment mechanism. And the look-then-move cycle shown in Fig. 2 is repeated until the final alignment error goes into the tolerance.

III. Coarse-to-Fine Two-Stage Alignment Control

Thus far, the notion of a coarse-to-fine control was raised as a trial to increase the servo bandwidth of a positioning system and popularly applied to dual-stage feed drive [9], macro-micro manipulator [10], data storage system [11], and so on. Commonly, they address how to coordinate the motion of two-stage positioning mechanisms which incorporate both coarse actuator and fine one. Recently, the coarse-fine control has been also applied to vision-based control systems closely concerned with the topic of this paper, for example, mobile robot localization [12], micrograsping [13], and microassembly [14]. In addition, the coarse-fine two-stage algorithm using an image pyramid was considered in the image processing area for fast template matching [15]. These approaches are fundamentally based on the data pruning technique to mitigate the computational burden coming from the large amount of vision data.

A. Observer-Based Motion Control

Due to the required time for the image processing necessary to get the alignment mark pose, most visual alignment systems in the display industry are running under the so-called look-then-move control. That is, visual sensing and actuation are performed in a sequential manner and the alignment stage keeps stationary during the image processing time. It causes each joint of the positioning mechanism to suffer from stick-slip motions due to the nonlinear friction around the zero velocity, which critically hinders microscopic fine control. In fact, as long as the look-then-move type visual servo is maintained, the slow update rate of vision data would result in repetitive vibratory motions in the vicinity of target point until the final misalignment gets smaller than the allotted tolerance.

To speed up the look-then-move recursion in Fig. 2, above all, a fast vision algorithm to extract the alignment marks should be given. However, if relying on a machine vision algorithm based on the general template matching, there is a certain limit in increasing the update rate because of the heavy computational time at the cost of generality. In this regard, we have suggested the geometric template matching (GTM) [3], [4] as an alignment mark specific algorithm which enables an order of reduction in the recognition time. It turns out to be a reasonable alternative for simple-shaped marks. However, when the data size of the raw images gets larger, even the specialized algorithm is not enough to achieve an update rate as fast as the capturing rate of the frame grabber.

Hence, we propose an observer-based coarse-to-fine visual alignment control system, in which both original fine images and reduced coarse ones are involved together. As the schematic in Fig. 3, it has a two-stage hierarchical structure, where the outer fine vision loop is to get exact misalignments for the original high resolution images and the inner coarse vision loop is aimed to update the misalignment state with a higher frequency as fast as the capturing rate of the CCD camera by incorporating reduced images with low resolution. The outer loop

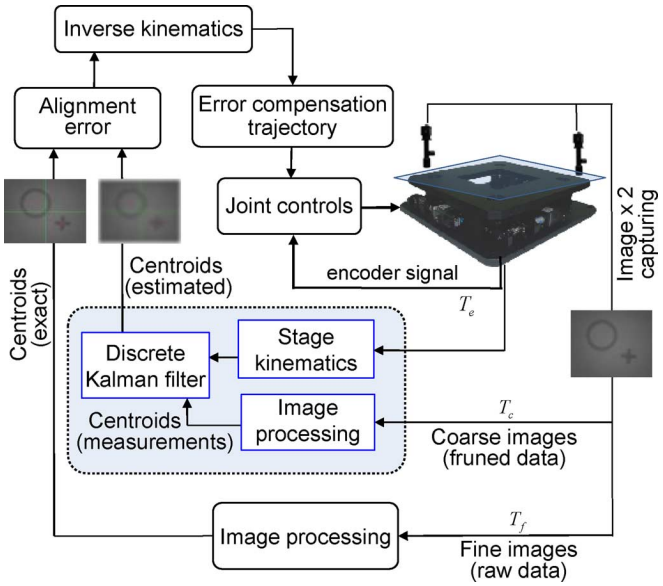

Fig. 3. Two-stage coarse-to-fine control loop for a visual alignment system.

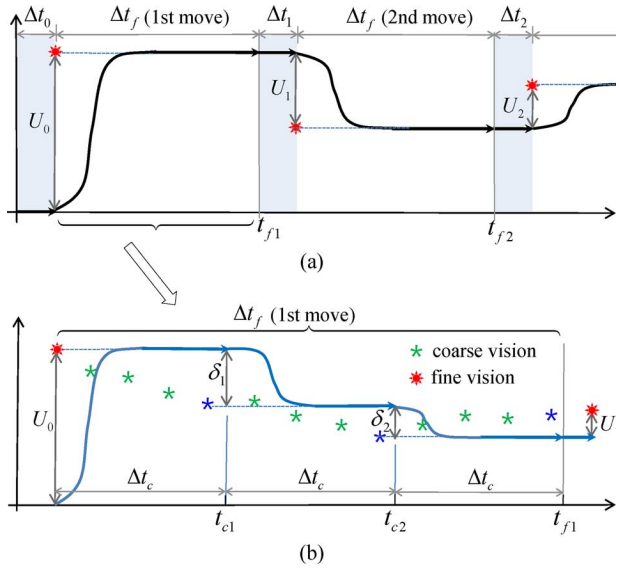

Fig. 4. Error compensation trajectory for a joint servo. $\Delta t_0 \sim \Delta t_n$: fine image processing times, U_0: initial misalignment, (δ_1, δ_2): intermittent error compensation distances by the coarse vision loop, U_1: misalignment after the first look-then-move, which is equal to the error compensation distance for the second look-then-move cycle. (a) Look-then-move trajectory (outer loop). (b) Look-and-move trajectory (inner loop).

in itself is actually equivalent to the conventional look-then-move alignment, while the inner loop can be called a look-and-move visual servo that the visual feedback and the control motion are performed simultaneously.

Although the use of pruned image data in the coarse vision loop enables a faster image processing, it brings about as less accurate results in localizing the alignment marks because of the decreased pixel resolution. In order to recover the degraded accuracy, the discrete Kalman filter [16]–[18] is adopted as a state observer to track the centroids of the moving alignment marks on the pattern mask in near real-time. In Fig. 3, the measurement update to the Kalman filter in terms of the coarse image processing is desired to be done almost as the image capturing speed and the period of the centroid prediction which requires the stage kinematics can be synchronized with the encoder signals of joint actuators. Then, the measurements and the predicted are fused in the recursion of Kalman filter to produce the outputs of enhanced accuracy.

Accordingly, a fast update of the misalignment state is achieved in the inner coarse vision loop and the outer fine vision loop is activated to confirm the misalignment every time the inner observer-based control loop is terminated. The look-and-move motion allows more frequent visual feedback and is advantageous in getting over frictional disturbances at each joint, which enables to outperform conventional alignment methods.

B. Error Compensation Trajectory Generation

For the in-plane mask-panel alignment, the alignment stages are required to have at least three active joints and most of them are with a parallel positioning mechanism (PPM) for high stiffness and high load capacity. In Fig. 3, the alignment error at the current time can be identified in task space according to the relationship in (1) and then resolved into joint servos in the forms of reference inputs for the misalignment compensation. The displacement of each joint can be determined by using the inverse kinematic solution for the given positioning mechanism. In the manufacturing process for flat panel displays, the mask-panel misalignment is at most hundreds of microns. This implies that a well-planned reference trajectory for the joint control loop is critical to get a good alignment performance.

In the two-stage alignment control structure, the error compensation trajectory of each active joint has also the two-stage hierarchy, as shown in Fig. 4, where the look-then- move trajectory of the outer loop embraces the look-and-move one of the inner loop. For example, when the initial misalignment (U_0) after reflecting the task space misalignment into the joint space is assigned to an active joint, the reference input for the joint servo can be determined as a smooth polynomial with proper rise time and termination time (t_{f1}). If the task space misalignment after the first look-then-move is beyond the set tolerance, the second cycle begins with a new decomposed misalignment (U_1). This process will be continued until the final misalignment meets the tolerance.

In fact, Fig. 4(a) alone corresponds to the case of conventional look-then-move alignment, where Δt_0 to Δt_n denote the image processing times to find the mask-panel misalignments and Δt_f represents joint control intervals. In contrast, in the observer-based coarse-to-fine alignment, the joint control trajectories, i.e., the move intervals in Fig. 4(a) which are solely based on the interoceptive joint sensor (motor encoder) are substituted into the look-and-move trajectories like Fig. 4(b) which also involves visual feedback, where the star marks at regular intervals represent the points at which the Kalman filter produces the estimates for the alignment mark pose in near real-time. The joint trajectory in Fig. 4 is just an example of $\Delta t_f = 3\Delta t_c$. In case of $\Delta t_f = \Delta t_c$, it means to give up the benefit of Kalman filter estimates.

The maximum update rate of the visual feedback is limited not only by the image processing time but also by the image capturing speed of camera. Hence, the realistic update rate of the alignment mark pose is at least tens of milliseconds. In this case, the direct vision-based control is not a good idea since it causes a large time delay in the feedback loop. As is shown in Fig. 4(b), the look-and-move trajectory consists of the consecutive individual polynomials which will be tracked by the encoder sensor-based joint control, where the error compensation distances are given by the visual feedback, the first one (U_0) by the fine vision loop and the other intermittent ones (δ_1, δ_2) by the coarse vision loop. As the coarse-to-fine visual feedback progresses, the misalignment at the end of each look-then-move cycle is expected to decrease more than when the single fine vision loop is applied. Then, the total alignment time until it meets the tolerance spec can be reduced as much by making the fine vision loop activated smaller times.

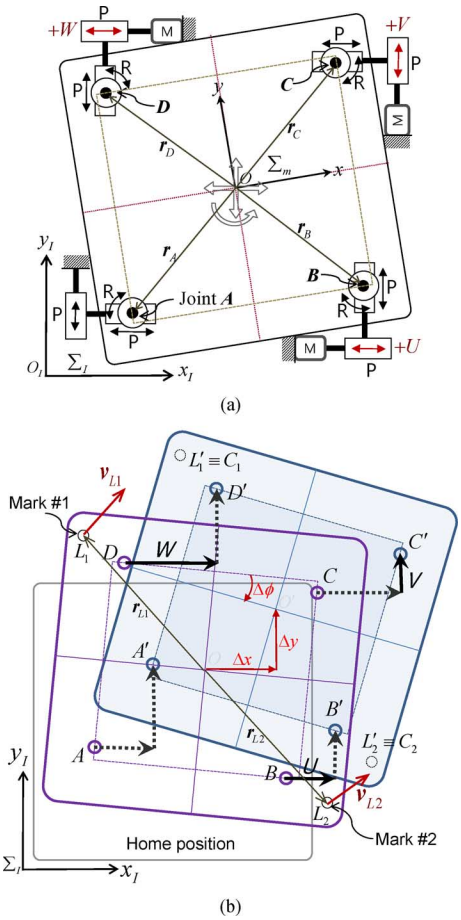

Fig. 5. The 4PPR alignment stage with three active joints. (a) Mobility of the 4PPR parallel mechanism. (b) Movement of the pattern mask (moving platform), solid lines: active joints, dotted lines: passive joints.

In the online generation of the error compensation trajectory, it belongs to the tuning problem to determine how frequent change of the polynomials according to the update of the misalignment state is desirable. In the microscopic low velocity motions of the alignment stage, if the period of polynomial is too short, specifically in the look-and-move trajectory, it tends to yield an adverse effect.

IV. Posture Estimation of Alignment Marks

In image processing problems, the Kalman filter [16] was popularly applied to overcome the difficulties coming from the bandwidth limit of vision sensors and the vast amount of image data, for example, face position tracking [17] and human body motion tracking [18]. In this section, it is demonstrated how to construct the Kalman filter to estimate the alignment mark postures, which is to increase the reliability of the coarse vision loop in Fig. 3.

A. Kinematics of the 4PPR Alignment Stage

As a positioning stage for the in-plane mask-panel alignment, we have adopted the 4PPR parallel mechanism with the schematic in Fig. 5(a), where all the limbs to support the moving platform consist of PPR (prismatic-prismatic-revolute) joints and the moving platform to determine the position of the pattern mask has 2T1R (two translational and one rotational) motion. First, we define the global frame Σ_I and the moving frame Σ_m attached to the platform. And it is assumed that the revolute joints (A, B, C, D) of the limbs are connected to the moving platform and the position vectors are given as $\boldsymbol{r}_A = (x_A, y_A)$, $\boldsymbol{r}_B = (x_B, y_B)$, $\boldsymbol{r}_C = (x_C, y_C)$, $\boldsymbol{r}_D = (x_D, y_D)$ with respect to Σ_m, Then, the inverse kinematic solution which relates between the in-plane misalignment $(\Delta x, \Delta y, \Delta\phi)$ and the corresponding displacements of the active prismatic joints (U, V, W) can be derived as follows [3]:

$$U = \Delta x + x_B(\cos\Delta\phi - 1) - y_B \sin\Delta\phi \tag{2a}$$

$$V = \Delta y + x_C \sin\Delta\phi + y_C(\cos\Delta\phi - 1) \tag{2b}$$

$$W = -\Delta x - x_D(\cos\Delta\phi - 1) + y_D \sin\Delta\phi. \tag{2c}$$

According to the controlled motion of the stage platform as denoted in Fig. 5(b), the two marks at L_1 and L_2 on the pattern mask are to be aligned to the glass panel marks at C_1, and C_2, respectively. Now, by differentiating (2) with respect to time, we have

$$\begin{bmatrix} \dot{U} \\ \dot{V} \\ \dot{W} \end{bmatrix} = \begin{bmatrix} 1 & 0 & -x_B\sin\Delta\phi - y_B\cos\Delta\phi \\ 0 & 1 & x_C\cos\Delta\phi - y_C\sin\Delta\phi \\ -1 & 0 & x_D\sin\Delta\phi + y_D\cos\Delta\phi \end{bmatrix} \begin{bmatrix} \dot{x}_o \\ \dot{y}_o \\ \dot{\phi} \end{bmatrix}$$
$$\Leftrightarrow \dot{\mathbf{q}}(t) = \mathbf{A}(\Delta\phi)\dot{\mathbf{p}}(t) \tag{3}$$

which defines the relationship between the velocity of active joints $\dot{\mathbf{q}}$ and the moving platform velocity $\dot{\mathbf{p}}$. Thus, the forward kinematic equation of the 4PPR alignment stage in velocity level is given by

$$\dot{\mathbf{p}}(t) = \mathbf{A}^{-1}\dot{\mathbf{q}}(t) = \mathbf{J}(\Delta\phi)\dot{\mathbf{q}}(t) \tag{4}$$

where the joint velocity can be obtained by using joint sensor (motor encoder) signals and the misalignment state in (1) must be given to compute the Jacobian matrix $\mathbf{J}(\Delta\phi)$ every sampling times.

Then, we have $\dot{\boldsymbol{r}}_o = (\dot{x}_o, \dot{y}_o, 0)$ the translational velocity of the origin and $\boldsymbol{\omega} = (0, 0, \dot{\phi})$ the angular velocity of the moving fame Σ_m with respect to Σ_I. As denoted in Fig. 5(b), we let $\boldsymbol{r}_{Li} = (x_{Li}, y_{Li}, 0)$ the position vector of the ith mark L_i $(i = 1,2)$ on the pattern mask from the origin of Σ_m to its centroid. Finally, the velocities of the alignment marks on the pattern mask can be determined as

$$\boldsymbol{v}_{Li}(t) = \dot{\boldsymbol{r}}_0 + \boldsymbol{\omega} \times \boldsymbol{r}_{Li} \Leftrightarrow \begin{bmatrix} v_{ix} \\ v_{iy} \end{bmatrix} = \begin{bmatrix} \dot{x}_0 - y_{Li}\dot{\phi} \\ \dot{y}_0 + x_{Li}\dot{\phi} \end{bmatrix} \ (i = 1, 2). \tag{5}$$

B. Measurement of the Alignment Mark Posture

Although a data pruning technique such as [15] and [19] could be considered for fast template matching, we are to apply the geometric template matching (GTM) [3], [4] which has been proved to be a reasonable approach for simple-shaped objects. Actually, the highlight of the image processing of the vision system in Fig. 2 lies in the pattern matching technique to recognize the alignment marks on the mask and panel.

The GTM is aimed to directly match predefined feature points in the designated matching area. As shown in Fig. 6, the geometric templates are defined as a collection of feature points enough to characterize the alignment mark shapes. Then, every pixel in the captured scene is scanned sequentially. For a specific pixel, if all the surrounding pixels at the same positions as the feature points have the same brightness in the binarized image, the central pixel will be registered as a candidate for the centroid. Because of the thickness of the marks, there could be a lot of candidates. For the cross marks, it needs to be rotated at each pixel. Then, the final centroidal position of each mark can be determined by averaging the positions of all the candidates.

C. Recursive Filtering Algorithm

In order to construct a motion model to describe the behavior of the alignment marks on the moving pattern mask, we can utilize the

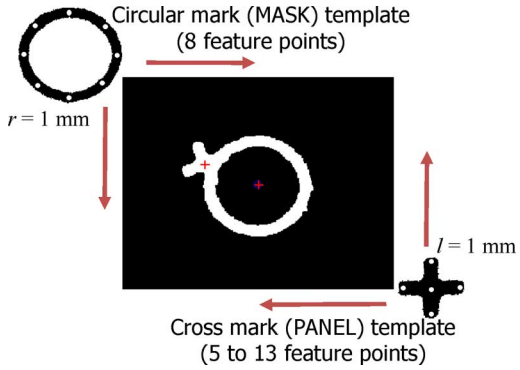

Fig. 6. Geometric template matching for the pattern recognition of the alignment marks.

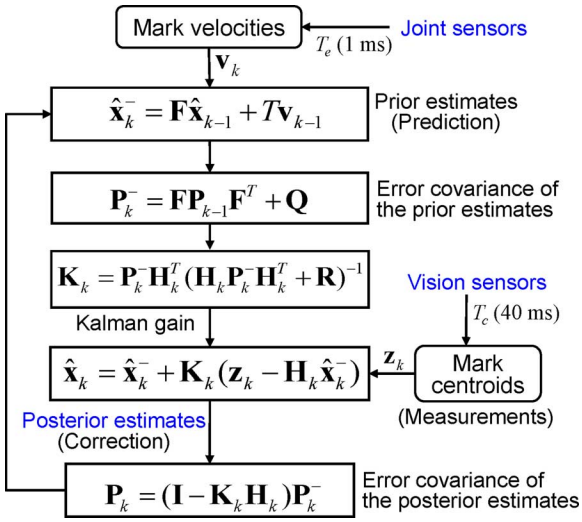

Fig. 7. Recursion of the discrete Kalman filter for the alignment mark posture estimation.

simple kinematic relationship: $\dot{\boldsymbol{r}}_{Li}(t) = \boldsymbol{v}_{Li}(t)$ between the centroidal position of the ith mark and its velocity. By assuming the backward approximation, we have

$$\begin{bmatrix} \dot{x}_{Li}(k) \\ \dot{y}_{Li}(k) \end{bmatrix} \approx \frac{1}{T}\begin{bmatrix} x_{Li}(k) - x_{Li}(k-1) \\ y_{Li}(k) - y_{Li}(k-1) \end{bmatrix} = \begin{bmatrix} v_{ix}(k) \\ v_{iy}(k) \end{bmatrix} (i = 1, 2). \tag{6}$$

Then, a discrete state-space equation for the in-plane position of the two marks can be written as

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{F}\mathbf{x}(k) + T\mathbf{v}(k) + \mathbf{w}(k) \\ \mathbf{x}(k) &= \begin{bmatrix} x_{L1} & y_{L1} & x_{L2} & y_{L2} \end{bmatrix}^T \in R^{4\times 1}, \\ &\quad \text{with} \\ \mathbf{v}(k) &= \begin{bmatrix} v_{1x} & v_{1y} & v_{2x} & v_{2x} \end{bmatrix}^T \in R^{4\times 1}, \mathbf{F} = \mathbf{I} \in R^{4\times 4} \end{aligned} \tag{7}$$

where $\mathbf{I}$ is the identity matrix, T the velocity update period, and $\mathbf{w}(k) \sim N(0, \mathbf{Q})$ denotes the white Gaussian process noise with zero mean and covariance $\mathbf{Q}$.

Second, as the measurement input to the Kalman filter, the centroids of the moving marks are extracted from the pruned data of coarse images. Since the centroidal positions are directly extracted through the image processing, the measurement equation can be written as

$$\mathbf{z}(k) = \mathbf{H}\mathbf{x}(k) + \mathbf{n}(k) \tag{8}$$

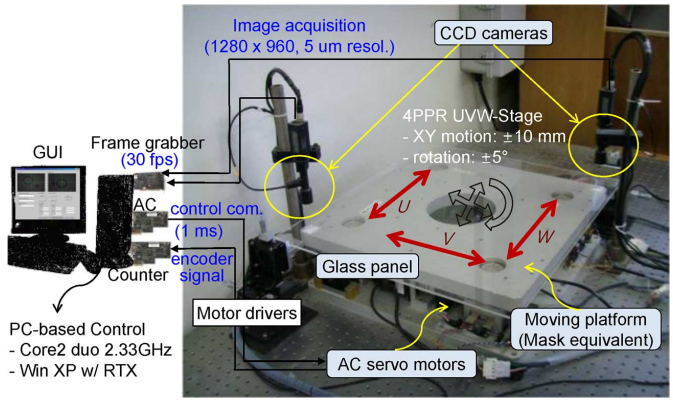

Fig. 8. Experimental setup for visual alignment control.

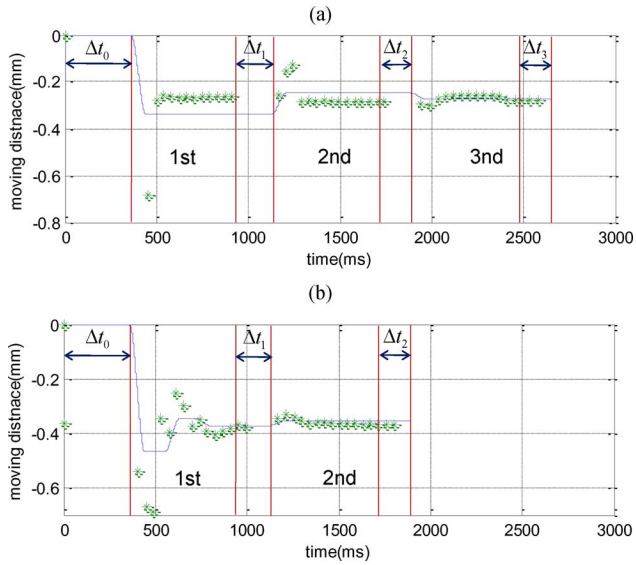

Fig. 9. Comparison of the error convergence time. (*) denotes the Kalman filter estimates resolved into U-axis of the alignment stage. (a) Look-then-move alignment only. (b) Coarse-to-fine two-stage alignment.

where $\mathbf{H} = \mathbf{I}$ and $\mathbf{n}(k) \sim N(0, \mathbf{R})$ is the zero-mean white Gaussian measurement noise.

Based on the motion model in (7) and the measurement one in (8), we have the recursive filtering algorithm in Fig. 7, which accepts the mark velocities at the rate of joint control by solving (5) and the mark centroids from the vision sensors with a proper update rate considering the capacity of the image processing algorithm. Finally, the posterior estimates after the measurement update corresponds to the output which will be used in the coarse-to-fine two-stage control.

V. Experimental Verification

The experimental setup in Fig. 8 has been constructed to evaluate the proposed visual alignment scheme, where the vision system is equipped with two CCD cameras of 1280×960 pixels, 5 μm resolution, 6.4×4.8 mm FOV, and illumination control. Also, the 4PPR positioning stage is working under PC-based joint control with 0.5 μm resolution encoder feedback. In order to implement the two-stage visual alignment, the size of the coarse image for the inner loop was determined as 320×240 pixels. And the Kalman filter produces the estimates every 40 ms by considering that the upper limit of the update rate of image data is 33.3 ms in the 30 fps frame grabber and the

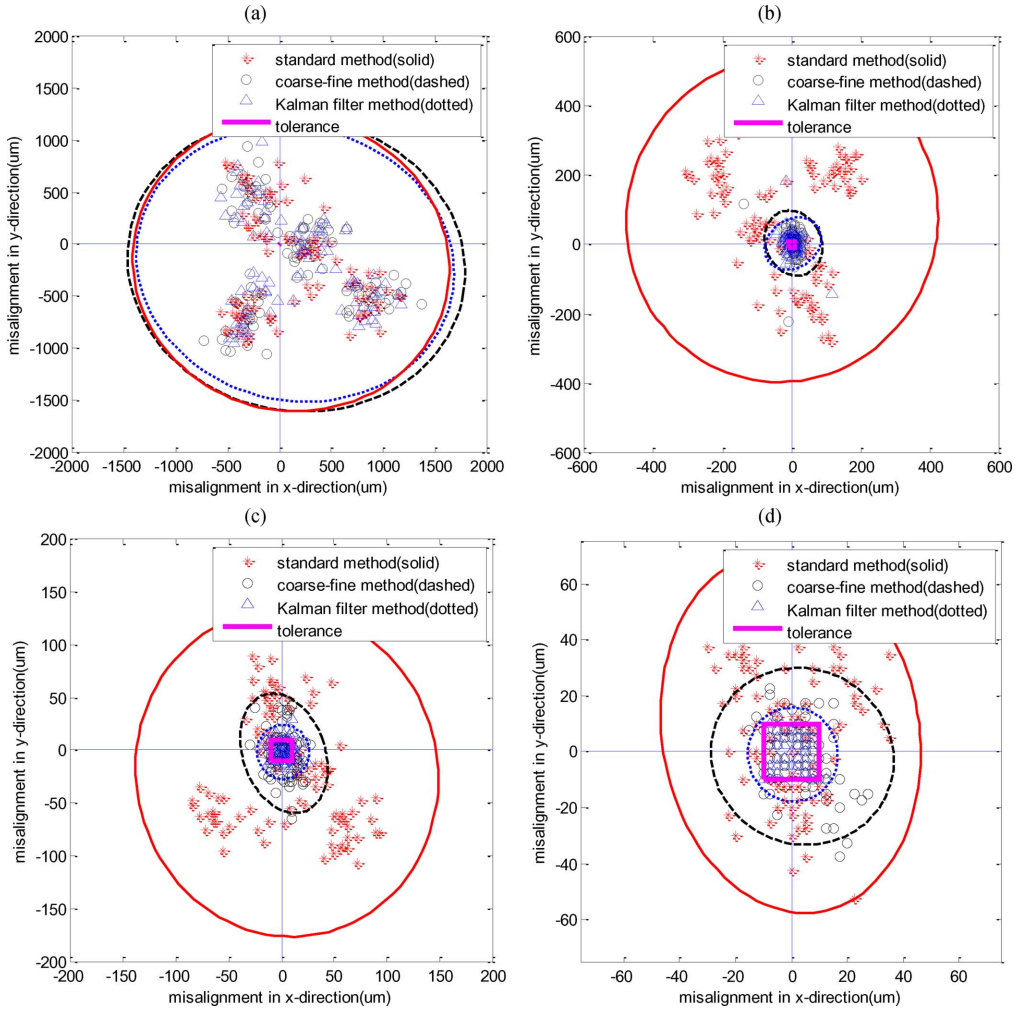

Fig. 10. Comparison of misalignment distributions (100 experiments), where the ellipses denote 3σ standard deviations in every direction. (a) Initial misalignments. (b) first look-then-move. (c) Second look-then-move. (d) Third look-then-move.

geometric template matching enables to get the mark centroids around 20–30 ms for the pruned image data of 320 × 240 pixels.

That is, the decomposed misalignment state denoted in Fig. 4(b) is renewed every 40 ms. After going through a performance tuning process, the polynomial trajectories to compensate for the intermittent misalignments in the observer-based control were determined to have the period of $\Delta t_c = 200$ ms with the rise time 80 ms and the settling time 120 ms. When it comes to the outer fine vision loop, it confirms the misalignment at the initial time and every three times recursion of the inner look-and-move trajectory. Then, the joint control time for the look-then-move cycle in Fig. 4(a) is $\Delta t_f = 3\Delta t_c$. That is, each look-then-move cycle includes three look-and- move periods. To achieve a good performance in the two-stage alignment control, it must be also guaranteed that the distributed joint servos of the alignment stage successfully follow the individual error compensation trajectories. We have applied the common PID control to each joint servo and the control gain sets were obtained through the tuning process.

As well, the Kalman filter parameters can be determined in such manner that the filtering outputs based on the coarse images are getting closer to the more exact results in terms of the fine images by adjusting the covariance $\mathbf{Q}$ and $\mathbf{R}$. It is reasonable to let the measurement noise covariance ($\mathbf{R}$) have the diagonal elements equal to the pixel resolution of input images. Then, the filtering performance and even more the whole alignment performance can be obtained by tuning the system noise covariance $\mathbf{Q}$. Finally, the following values were applied in getting the experimental results below:

$$\mathbf{Q} = 10\mathbf{I} \in R^{4\times4},\ \mathbf{R} = 0.02\mathbf{I} \in R^{4\times4}\ (\mathrm{mm}). \tag{9}$$

First, Fig. 9 is a specific example to illustrate how the two-stage control decreases the total alignment time, where the solid lines represent the traces of the U-axis motion under encoder-based joint control and the star marks the Kalman filer estimates. When only the look-then-move alignment (i.e., the fine-vision loop in Fig. 3) was applied, the Kalman filter outputs were not involved in the control loop and the alignment process ended at the third cycle. However, in case of the two-stage alignment, the estimation results were reflected to change the joint reference trajectory according to the strategy suggested in Fig. 4(b) and the alignment process was finished in the second cycle, which indicates as much reduction of the alignment time.

Fig. 9 is just one of the results which could happen in the mask-panel alignment. Depending on the initial misalignment and other uncertainty effects, the results will become different and sometimes the conventional method could perform better. Hence, we have repeated the experiment 100 times in order to evaluate the alignment performance in a statistical manner. The error budgets in Fig. 10 shows how the misalignment distributions converge into the tolerance box according as

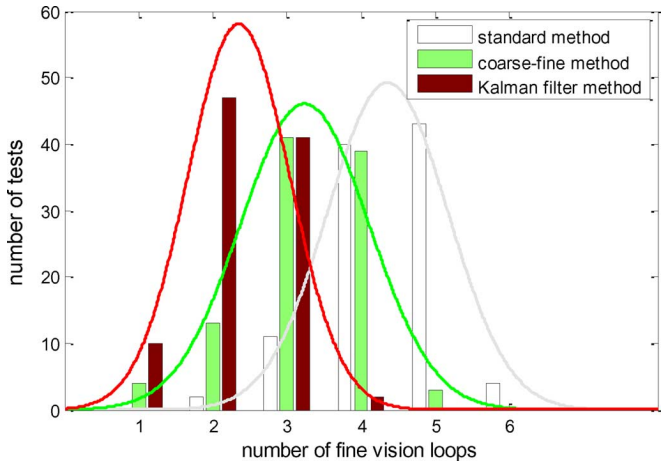

Fig. 11. Histogram of the number of fine vision loop activations and the transfer of the normal distributions (100 experiments).

TABLE I
NUMBER OF FINE VISION LOOP ACTIVATIONS

	Mean	Standard deviation
standard method	4.36	0.8105
coarse-fine method	3.24	0.8660
Kalman filter method	2.35	0.6872

the look-then-move cycle is progressed, where: i) "standard method"' means the conventional look-then-move alignment which are widely adopted in the display industry; ii) "coarse-fine method" is the case when the two-stage control is applied but the alignment mark poses are directly given by the measurements from the reduced images instead of the Kalman filter outputs; and iii) "Kalman filter method" denotes the two-stage alignment to utilize the filtering outputs. As shown, the two-stage control achieves much faster convergence to the tolerance compared with the conventional method. Also, it is certain that the Kalman filter contributes to accelerating the speed of convergence.

In Fig. 10, the initial misalignments were generated by the home positioning actions of the alignment stage and the group of ellipses represents the 3σ band of misalignment in all directions. In actual display manufacturing lines, the alignment stage carries much heavier weights and the total alignment time roughly ranges from 15 to 25 s depending on the panel size, but it includes the vertical attachment time of mask and panel before image processing begins. Also, the alignment accuracy specification is usually given as two times of the CCD pixel resolution.

The histogram in Fig. 11 statistically evaluates the number of fine vision loop activations until the alignment process is terminated depending on the alignment methods, while the means and the standard deviations are listed in Table I. As a consequence, about 46% improvement was accomplished by the proposed control. Thus, we can expect as much reduction of alignment time. Finally, the normal distribution curves in Fig. 11 indicate that the dispersion of the number of fine vision loop repetitions gets narrower in the two-stage alignment, which means that the proposed control tends to make the alignment time more consistent.

VI. CONCLUDING REMARKS

This paper is providing a new control framework for the visual alignment systems and the engineering process to organize all the algorithms required for the real implementation. Considering the vision system to get the alignment mark postures, high-speed cameras are meaningless as far as the image processing time is over the capturing period. In general, current 30 fps cameras are good enough for the visual alignment applications. If a high-performance image processing system with the maximum computing power is available, the update rate for the centroid measurements in the Kalman filter can be increased as much. Then, the look-and-move period in Fig. 4(b) will get faster and the coarse-to-fine alignment would be more promising. Another consideration can be taken into the design of joint servos for the alignment stage to follow the error compensation trajectory. Although we have applied PID control for each active joint, a model-based control design to reflect the friction characteristics of the moving axes would be more effective in the precision control.

REFERENCES

[1] A. K. Kanjilal, "Automatic mask alignment without a microscope," *IEEE Trans. Instrumen. Meas.*, vol. 44, no. 3, pp. 806–809, Jun. 1995.
[2] H. T. Kim, C. S. Song, and H. J. Yang, "2-step algorithm of automatic alignment in wafer dicing process," *Microelectron. Reliab.*, vol. 44, pp. 1165–1179, 2004.
[3] S. J. Kwon and J. Hwang, "Kinematics, pattern recognition, and motion control of mask-panel alignment system," *Control Eng. Practice*, vol. 19, pp. 883–892, 2011.
[4] C. S. Park and S. J. Kwon, "An efficient vision algorithm for fast and fine mask-panel alignment," in *Proc. SICE-ICASE Int. Joint Conf.*, Oct. 2006, pp. 1461–1465.
[5] S. J. Kwon and H. Jeong, "Observer based fine motion control of autonomous visual alignment systems," in *Proc. 2009 IEEE/ASME Int. Conf. Adv. Intell. Mechatron.*, Singapore, Jul. 14–17, 2009, pp. 1822–1827.
[6] L. Alvarez, J. Yi, R. Horowitz, and L. Olmos, "Dynamic friction model-based tire-road friction estimation and emergency braking control," *ASME J. Dynamic Syst., Meas., Control*, vol. 127, pp. 22–32, Mar. 2005.
[7] D. Chwa and J. Y. Choi, "Observer-based control for tail-controlled skid-to-turn missiles using a parametric affine model," *IEEE Trans. Control Syst. Technol.*, vol. 12, no. 1, pp. 167–175, Jan. 2004.
[8] A. Swarnakar, H. J. Marquez, and T. Chen, "A new scheme on robust observer-based control design for interconnected systems with application to an industrial utility boiler," *IEEE Trans. Control Syst.s Technol.*, vol. 16, no. 3, pp. 539–548, May 2008.
[9] A. T. Elfizy, G. M. Bone, and M. A. Elbestawi, "Design and control of a dual-stage feed drive," *Int. J. Mach. Tools Manuf.*, vol. 45, pp. 153–165, 2005.
[10] J. B. Morrell and J. K. Salisbury, "Parallel-coupled micro-macro actuators," *Int. J. Robot. Res.*, vol. 17, no. 7, pp. 773–791, 1998.
[11] T. Semba, T. Hirano, J. Hong, and L. Fan, "Dual-stage servo controller for HDD using MEMS microactuator," *IEEE Trans. Magn.*, vol. 35, no. 5, 1999.
[12] J. Wang, H. Zha, and R. Cipolla, "Coarse-to-fine vision-based localization by indexing scale-Invariant features," *IEEE Trans. Syst., Man, Cybern.-Part B: Cybern.*, vol. 36, no. 2, pp. 413–422, Apr. 2006.
[13] L. Ren, L. Wang, J. K. Mills, and D. Sun, "Vision-based 2-D automatic micrograsping using coarse-to-fine grasping strategy," *IEEE Trans. Ind. Electron.*, vol. 55, no. 9, pp. 3324–3331, Sep. 2008.
[14] S. J. Ralis, B. Vikramaditya, and B. J. Nelson, "Micropositioning of a weakly calibrated microassembly system using coarse-to-fine visual servoing strategies," *IEEE Trans. Electron. Packag. Manuf.*, vol. 23, no. 2, pp. 123–131, Apr. 2000.
[15] M.-S. Choi and W.-K. Kim, "A novel two stage template matching method for rotation and illumination invariance," *Pattern Recognit.*, vol. 35, pp. 119–129, 2002.
[16] M. S. Grewal and A. P. Andrews, *Kalman Filtering: Theory and Practice Using Matlab*. New York: Wiley, 2001.
[17] R. Qian, M. Sezan, and K. Matthews, "A robust real-time face tracking algorithm," in *Proc. IEEE Int Conf. Image Process. (ICIP 98)*, Oct. 1998, pp. 131–135.
[18] D. S. Jang, S. W. Jang, and H. I. Choi, "2D human body tracking with structural Kalman filter," *Pattern Recognit.*, vol. 35, pp. 2041–2049, 2002.
[19] M. Gharavi-Alkhansari, "A fast globally optimal algorithm for template matching using low-resolution pruning," *IEEE Trans. Image Process.*, vol. 10, no. 4, pp. 526–533, Apr. 2001.