

# Model Adaptation and Posture Estimation of Moving Articulated Object Using Monocular Camera

N. Shimada<sup>†</sup>, Y. Shirai<sup>†</sup> and Y. Kuno<sup>†,\*</sup>

E-mail: shimada@mech.eng.osaka-u.ac.jp

<sup>†</sup> Dept. of Computer-Controlled  
Mechanical Systems, Osaka University  
2-1 Yamadaoka, Suita,  
Osaka, 565-0871, Japan

\* Dept. of Information and Computer  
Sciences, Saitama University  
255, Shimo-okubo, Urawa, Saitama  
338-8570, Japan

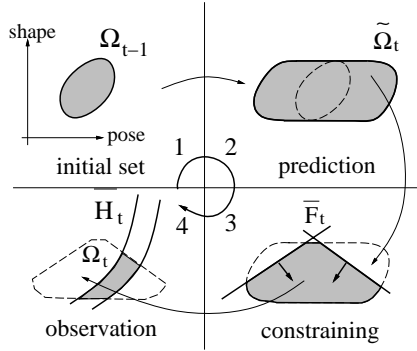
## Abstract

*This paper presents a method of estimating both 3-D shapes and moving poses of an articulated object from a monocular image sequence. Instead of using direct depth data, prior loose knowledge about the object, such as possible ranges of joint angles, lengths or widths of parts, and some relationships between them, are referred as system constraints. This paper first points out that the estimate by Kalman filter essentially converge to a wrong state for non-linear unobservable systems. Thus the paper proposes an alternative method based on a set-membership-based estimation including dynamics. The method limits the depth ambiguity by considering loose constraint knowledge represented as inequalities and provides the shape recovery of articulated objects. Effectiveness of the framework is shown by experiments.*

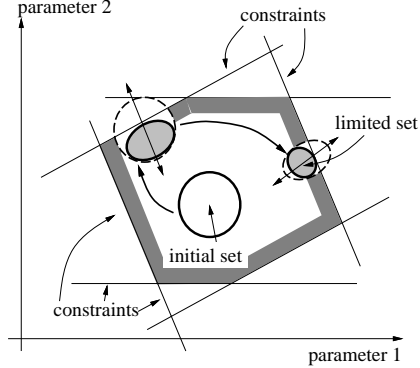
## 1 Introduction

In general, the 3-D shape of a non-rigid object cannot be recovered with only one camera even if an image sequence is given. For a certain class of objects like human bodies, however, depth information can be recovered from a monocular camera data if adequate prior knowledge about the shape and structure is available. For example, we can build a structure model of a human body by assuming that its joint motions are related with each other and there are rough correlations between the sizes of body parts. Given such constraints, depth estimation can be broken down into a least squares problem [1][2]. Particularly, Kalman filter and its variation for non-linear systems (Extended Kalman filter) are considered to give estimates with good accuracy and popularly applied to estimate an object's shape or pose parameters. If the constraints are rough to be represented by inequalities, this method cannot be applied because the original EKF cannot deal with the inequalities. Shimada et.al [3] recovered the joint angles and the finger lengths from a monocular image sequence using an EKF variation handling the inequalities by distribution truncation.

In addition, Kalman filter has an essential problem. Although it can give correct estimates only for linear systems and non-linear observable systems, its estimate converges to a wrong state due to approximation error in linearization for non-linear unobservable systems. As an alternative way, the paper proposes an estimation method based on *set-membership estimation*. In EKF, the ambiguity of the estimate is described by the estimate covariance. In the set-membership methods, a possible parameter set described by its boundary is used for representing the estimate ambiguity, which is represented by



**Fig.1.** Incremental update of parameter set



**Fig.2.** Possibility limitation by constraint knowledge

the estimate covariance in EKF. In this paper, the combination of the ellipsoidal description[4][5] and a rectangular description is proposed. In order to estimate time-varying parameters, the proposed framework introduces an updating scheme by dynamics. Since this method does not calculate probabilistic integration but an intersection set of possible parameters, it can provide more accurate estimation than EKF because it can avoid the accumulation of linearization error. For this ability, loose constraint knowledge such as possible ranges of lengths, widths or joint angles are available to estimate the shape and pose of articulated objects. While the set-membership-based approach has these merits, it is well-known that it has a drawback of the weakness against outliers. Here we concentrate on describing the integration scheme at each time-frame supposing the existence of an outlier rejection method.

In the following sections, the basic idea is first explained and the details are described later. Finally effectiveness of the method is shown by estimation results for an articulated object.

## 2 Basic Idea

Although monocular imaging systems have unobservability of depth in shape estimation of non-rigid objects, they can estimate depth information if prior knowledge is additionally available. For example, we can recover a human body by approximating it as an articulated object which consists of rigid body parts linked each other and by assuming the following constraints.

- (a) Shape parameters (lengths and widths) are constant over an image sequence.
- (b) Pose parameters (joint angles and orientation) change continuously.
- (c) Each parameter is within a certain range and has relations with other parameters.

Note that these constraints include not only equations but also loose constraints described as inequalities.

Fig. 1 summarizes the basic idea of parameter estimation with these constraints under “unobservable” systems<sup>1</sup>. A vector space of parameters describing shape and pose is first considered and an initial possible parameter set  $\Omega_{t-1}$  is supposed. Then a predicted

<sup>1</sup> The term “unobservability” is used in the control theory. It means that there are states which cannot be discriminated by any observation sequences.

set  $\tilde{\Omega}_t$  is generated by shifting and diffusing the initial set based on the object's dynamics and its noise. When an observation data is obtained from each image, a parameter set  $\tilde{H}$  satisfying the observation is considered. Another set  $\tilde{F}$  satisfying the constraint knowledge is considered as well. Then the updated possible parameter set  $\Omega_t$  is obtained as an intersection of the three parameter sets. This process is repeated for each image observation. The important point is that the constraint knowledge has an effect to limit unobservable modes as illustrated in Fig. 2. Until the parameter set intersects with any boundary of the constraints, the system is still unobservable and the ambiguity remains. Once it intersects with a constraint boundary, however, the possible parameter set is limited to only the inside of the boundary. If any parameters are supposed to be constant, such as lengths or widths of rigid parts, the ranges of these parameters should monotonously decrease. As the result of iterative intersection with boundaries during sequential observations, ranges of any parameters including unobservable modes can be limited.

### 3 Limit of EKF Estimation

#### 3.1 Non-linear Unobservable System

The idea described in the previous section is naturally implemented in a state estimation framework of Kalman filter or its non-linear version (Extended Kalman filter, EKF). Shimada et al. utilized EKF to implement the idea [3].

Suppose the transition and observation formulas of a system represented as

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{u}_t \quad (1)$$

$$\mathbf{y}_t = \mathbf{h}(\mathbf{x}_t) + \mathbf{w}_t \quad (2)$$

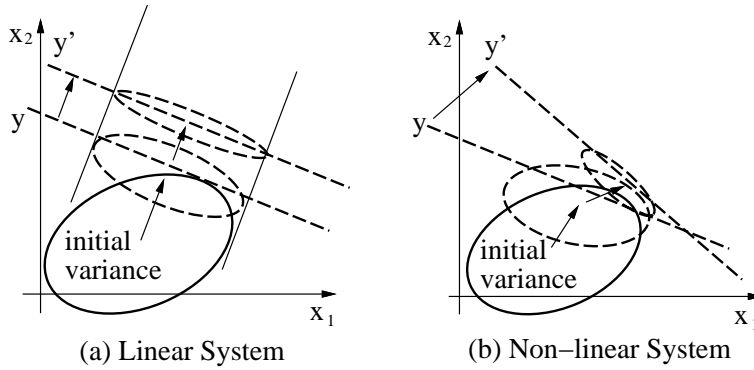
where  $\mathbf{y}_t$  is an observation vector.  $\mathbf{u}_t$  and  $\mathbf{w}_t$  are white noises with zero mean and variances  $\mathbf{U}$ ,  $\mathbf{W}$ . The components of  $\mathbf{U}$  for time-invariant shape parameters are zeros. If the observation function  $\mathbf{h}(\cdot)$  is non-linear, the current state  $\hat{\mathbf{x}}_t$  and variance  $\mathbf{P}_t$  are approximately estimated by EKF:

$$\hat{\mathbf{x}}_t = \tilde{\mathbf{x}}_t + \mathbf{K}_t\{\mathbf{y}_t - \mathbf{h}(\tilde{\mathbf{x}}_t)\} \quad (3)$$

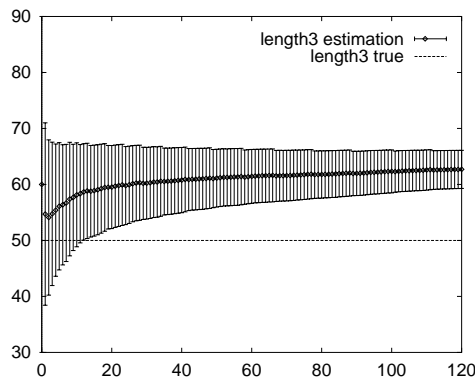
$$\mathbf{P}_t = (\mathbf{I} - \mathbf{K}_t\mathbf{h}'(\tilde{\mathbf{x}}_t))(\mathbf{A}\mathbf{P}_{t-1}\mathbf{A}^T + \mathbf{B}\mathbf{U}\mathbf{B}^T) \quad (4)$$

where  $\mathbf{K}_t$  is a Kalman gain matrix and  $\tilde{\mathbf{x}}_t = \mathbf{A}\hat{\mathbf{x}}_{t-1}$ .

EKF is recognized as the estimator to achieve good accuracy. It is correct for linear KF, but not for EKF. Actually the EKF estimate for non-linear unobservable systems essentially converges to a wrong state under a certain condition. Fig. 3 illustrates the reason. In linear systems, the direction of a hyper-plane satisfying a certain observation  $\mathbf{y}$  is invariant. Therefore the co-variance ellipsoid decreases only in the observable directions and ambiguity of unobservable modes is properly preserved. In contrast, the direction of the hyper-plane depends on the linearizing point in non-linear systems. Therefore the variance ellipsoid is destined to shrink in any directions even if the system is actually unobservable. When all parameters are time-varying, this is not a serious problem because the erroneous shrinkage is swallowed up by the variance increase in any directions at each time step due to the system transition noise. However, it is fatal for systems including time-invariant parameters such as a length or width of rigid objects. When constant observations only perturbed by noise are obtained, the EKF estimate converges into a wrong state as shown in Fig. 4. The vertical bars mean the range of 2 standard deviation. Once the wrong conversion occurs, the estimate is difficult to modify due to too small variance even if observations containing new information are obtained. Then tracking fails as a result. Fig. 5 shows such a failure example for a 3-link arm in 2-D in which the black circle denotes the fixed origin of the arm and the open circles corresponds to the link-joints or the most proximal tip of the arm.



**Fig. 3.** Linearization error for non-linear unobservable systems



**Fig. 4.** Wrong conversion

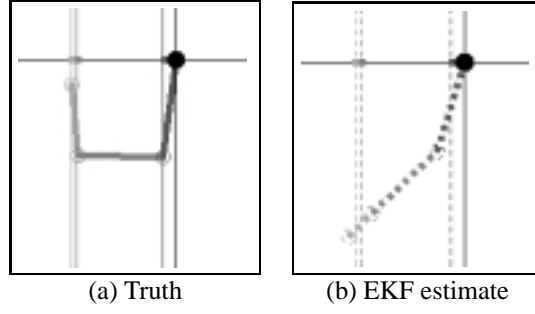
### 3.2 Inequality Constraints

The constraints (b) and (c) introduced in section 2 include loose constraints represented as inequalities. Although the equation constraints can be treated as an observation with a zero variance in the EKF framework, inequalities cannot be dealt in such a simple way. Hel-Or et al.[6] modified the inequality constraints to quadratic equations with slack variables to linearize them. The linearized constraints, however, are quite different from the original ones. Another way is to introduce the constraints as an initial distribution. It is also inappropriate because the effect of the initial distribution decreases at every frame.

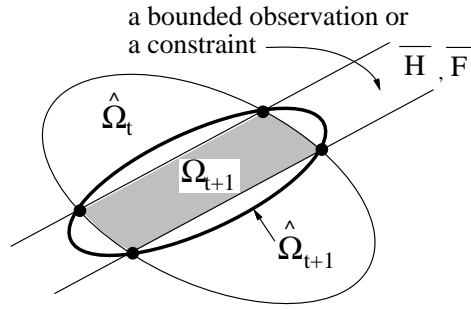
## 4 Possibility Reduction based on Set-membership

### 4.1 Ellipsoidal Boundary Description

The set-membership methods are known as a way to estimate an unknown system parameters from a sequence of input signals and the corresponding output of the system under bounded signal errors. The methods represent a possible parameter set as intervals of each parameter[7], a polygon[8] or an ellipsoid[4] in a high-dimensional vector



**Fig. 5.** Wrong Estimate by EKF



**Fig. 6.** Update of ellipsoidal set-membership

space. Fogel et al.[5] formulated a set updating with an ellipsoidal initial set and a linear observation with bounded noise.

Suppose an initial ellipsoid  $\hat{\Omega}_t$  described as

$$(\mathbf{x}_t - \mathbf{q}_t)^T \mathbf{P}_t^{-1} (\mathbf{x}_t - \mathbf{q}_t) \leq 1 \quad (5)$$

and a parameter set  $\bar{H}$  satisfying a noise-bounded observation as

$$|\mathbf{m}_i^T \mathbf{x}_t - \rho_i| \leq c_i. \quad (6)$$

The latter means the region between two hyper-planes. A parameter set  $\bar{F}$  satisfying an inequality constraint is described as well. A certain ellipsoid  $\hat{\Omega}_{t+1}$  covering whole the intersection set  $\Omega_{t+1}$  between  $\hat{\Omega}_t$  and  $\bar{H}(\bar{F})$  (see Fig. 6) can be represented with  $\lambda \geq 0$  as

$$(\mathbf{x}_t - \mathbf{q}_t)^T \mathbf{P}_t^{-1} (\mathbf{x}_t - \mathbf{q}_t) + \lambda (\mathbf{m}_i^T \mathbf{x}_t - \rho_i)^2 \leq 1 + \lambda c_i^2. \quad (7)$$

which can be reformed to the form of Eq. 5. Using a criterion minimizing  $\text{tr} \mathbf{P}_{t+1}$ ,  $\lambda$  achieving the smallest ellipsoid is the real root of

$$\beta_1 \lambda^3 + \beta_2 \lambda^2 + \beta_3 \lambda + \beta_4 = 0 \quad (8)$$

where

$$\beta_1 = G^2 c_i^2 (\mu G - \gamma),$$

$$\begin{aligned}
\beta_2 &= 3Gc_i^2(\mu G - \gamma), \\
\beta_3 &= G\mu(3c_i^2 - \varepsilon^2) - \gamma\{2(c_i^2 - \varepsilon) + G\}, \\
\beta_4 &= \mu(c_i^2 - \varepsilon^2) - \gamma, \\
\mu &= \text{tr}\mathbf{P}_t, \\
\gamma &= \mathbf{m}_i^T \mathbf{P}_t^2 \mathbf{m}_i, \\
G &= \mathbf{m}_i^T \mathbf{P}_t \mathbf{m}_i, \\
\varepsilon &= \rho_i - \mathbf{m}_i^T \mathbf{q}_t.
\end{aligned} \tag{9}$$

It can be proved that Eq. 8 always has one real root and two imaginary roots. In addition, if the real root is negative, the minimal  $\text{tr}\mathbf{P}_{t+1}$  in the domain of  $\lambda \geq 0$  is achieved at  $\lambda = 0$  (the details of the proof are described in [5]). Finally, the updated ellipsoid is described in the form of Eq. 5 by

$$\begin{aligned}
\mathbf{q}_{t+1} &= \mathbf{q}_t + \lambda \mathbf{Y} \mathbf{m}_i \varepsilon, \\
\mathbf{P}_{t+1} &= z \mathbf{Y}
\end{aligned} \tag{10}$$

where

$$\begin{aligned}
\mathbf{Y} &= \mathbf{P}_t - \lambda \frac{\mathbf{P}_t \mathbf{m}_i \mathbf{m}_i^T \mathbf{P}_t^T}{1 + \lambda G}, \\
z &= 1 + \lambda c_i^2 - \frac{\lambda \varepsilon^2}{1 + \lambda G}.
\end{aligned} \tag{11}$$

Since this method does not calculate a probability variance but a boundary of the intersection, the parameter set should keep its size even if almost the same observations are repeatedly obtained.

## 4.2 Extension to Dynamic Systems

Most conventional set-membership methods contained no dynamic mechanism[9] or treated dynamics for low dimensional systems using a simple interval description[7]. While update by dynamics can be easily introduced into such interval descriptions, it is unable to represent correlations between each parameter. In addition, the set description is inaccurate. The ellipsoidal set-membership can be easily updated in high dimensional systems and has the ability to describe correlations. Thus we introduce a dynamics and prediction phase into the ellipsoid-based method.

Suppose a dynamics of the system is given as Eq. 1 and the system noise  $\mathbf{u}_t$  is bounded as

$$\mathbf{u}_t^T \mathbf{U}^{-1} \mathbf{u}_t \leq 1. \tag{12}$$

In the  $(\mathbf{x}_t, \mathbf{u}_t)$  space, a certain ellipsoid including all of the parameters satisfying Eqs. 5 and 12 can be represented as

$$(\mathbf{x}_t - \mathbf{q}_t)^T \mathbf{P}_t^{-1} (\mathbf{x}_t - \mathbf{q}_t) + \lambda \mathbf{u}_t^T \mathbf{U}^{-1} \mathbf{u}_t \leq 1 + \lambda \tag{13}$$

where  $\lambda \geq 0$ . Combining this and Eq.1,

$$\mathbf{q}_{t+1} = \mathbf{A} \mathbf{q}_t \tag{14}$$

$$\mathbf{P}_{t+1} = (1 + \lambda) \mathbf{A} \mathbf{P}_t \mathbf{A}^T + (1 + \frac{1}{\lambda}) \mathbf{B} \mathbf{U} \mathbf{B}^T \tag{15}$$

is obtained. See Appendix for details. Here, using a criterion which minimizing  $\text{tr}\mathbf{P}_{t+1}$ , the smallest ellipsoid is given by

$$\lambda = \sqrt{\frac{\text{tr}(\mathbf{B}\mathbf{U}\mathbf{B}^T)}{\text{tr}(\mathbf{A}\mathbf{P}_t\mathbf{A}^T)}}. \quad (16)$$

After this predictional update, the observations and constraints are integrated in the way described in section 4.1.

### 4.3 Maximum and Minimum Bounds

In updating by Eq. 11, the parameter set can be represented as an small ellipsoid if  $c_i$  is small enough compared to the size of the original ellipsoid. Otherwise, the updated ellipsoid tends to include a large region outside the true intersection  $\Omega_t$ . Fogel et al.[5] added a pre-process: when any of the two hyper-plane boundaries of Eq. 6 does not intersect with the ellipsoid to update,  $\rho_i$  and  $c_i$  are modified so that the non-intersecting hyper-plane is tangent to the ellipsoid. However, if the ellipsoid to update highly sticks out of the hyper-plane boundary, the new ellipsoid simply updated by Eq. 11 is almost the same as the original ellipsoid.

In our method, the maximum and minimum values of each component of  $\mathbf{x}$  are calculated by controlling  $\lambda$  in Eq. 7. The maximum and minimum in direction  $\mathbf{n}$  ( $|\mathbf{n}| = 1$ ) are obtained by substituting a new criterion minimizing  $\mathbf{n}^T\mathbf{P}_{t+1}\mathbf{n}$  for minimizing  $\text{tr}\mathbf{P}_{t+1}$ .  $\lambda$  achieving that criterion is obtained by solving Eq. 8 where

$$\begin{aligned} \beta_1 &= G^2 c_i (GK - L^2) \\ \beta_2 &= 3Gc_i^2 (GK - L^2) \\ \beta_3 &= GK(3c_i^2 - e^2) + L^2(2e^2 - 2c_i^2 - G) \\ \beta_4 &= (c_i^2 - e^2)K - L^2 \\ K &= \mathbf{n}^T \mathbf{P}_t \mathbf{n} \\ L &= \mathbf{m}_i^T \mathbf{P}_t \mathbf{m}_i. \end{aligned} \quad (17)$$

The maximum  $x_{max}^{(t)}$  and minimum  $x_{min}^{(t)}$  are obtained by

$$x_{min}(\mathbf{n}) = \mathbf{n}^T \mathbf{s} - \sqrt{\mathbf{n}^T \mathbf{Q} \mathbf{n}} \quad (18)$$

$$x_{max}(\mathbf{n}) = \mathbf{n}^T \mathbf{s} + \sqrt{\mathbf{n}^T \mathbf{Q} \mathbf{n}} \quad (19)$$

$$\mathbf{s} = \mathbf{q}_t + \lambda \mathbf{Y} \mathbf{m}_i \varepsilon \quad (20)$$

$$\mathbf{Q} = z \mathbf{Y} \quad (21)$$

where  $z$  and  $\mathbf{Y}$  are given by Eq. 11. If  $x_{max}$  and  $x_{min}$  for any  $\mathbf{n}$  can be calculated, a convex hull of the true intersection set  $\Omega_{t+1}$  is obtained. Since systems for articulated objects have much high dimensions, each coordinate axis of the parameter space is used as a representative. Whenever an observation is obtained and the ellipsoid is updated in time  $t$ ,  $x_{min}^{(t)}$  and  $x_{max}^{(t)}$  for each parameter are calculated and then  $\max(x_{min}^{(t)}, x_{min}^{(t-1)})$  and  $\min(x_{max}^{(t)}, x_{max}^{(t-1)})$  are preserved. At the final stage of each update, a constraint in the form of Eq. 6 is made from  $x_{min}^{(t)}$  and  $x_{max}^{(t)}$  and the ellipsoid is updated by the constraint in the way of Eq. 11.

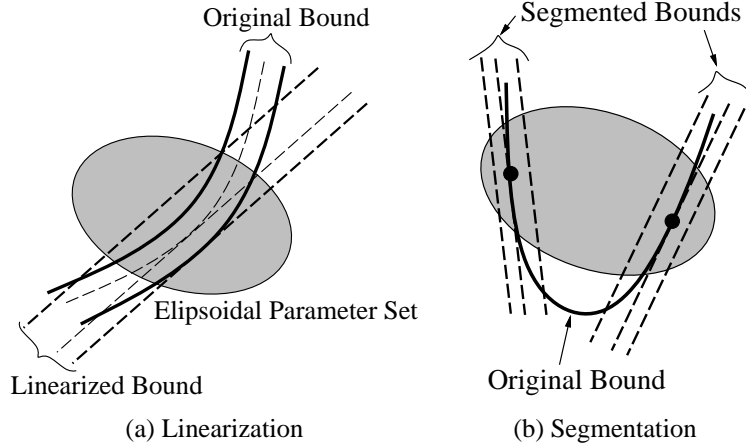


Fig. 7. Compensation for non-linearity of bound

#### 4.4 Compensation for Non-linearity of Bounds

In the previous section, we summarize our set-membership-based estimation supposing the observation and constraint formulas are linear as Eq. 6. Since they are actually highly non-linear like Eq. 2, they need to be approximated as a linear form:

$$\mathbf{y}_t = \mathbf{h}'(\tilde{\mathbf{x}}_t)(\mathbf{x}_t - \tilde{\mathbf{x}}_t) + \mathbf{h}(\tilde{\mathbf{x}}_t) + \mathbf{w}_t. \quad (22)$$

Since this approximation includes linearization error,  $\rho$  and  $c$  should be determined so that the parameter set satisfying Eq. 6 includes one satisfying both Eq. 22 and noise bound

$$\mathbf{w}_t \mathbf{W}^{-1} \mathbf{w}_t \leq 1. \quad (23)$$

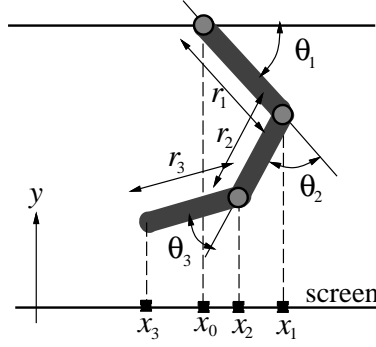
First a linearizing point is found by iterative solving of Eq. 22 with an initial solution  $\tilde{\mathbf{x}}_t$ . Then Eq. 2 is linearized and decomposed into each component  $y_i$ . The decomposed  $\mathbf{h}'_i$  is determined as  $\mathbf{m}_i$  in Eq. 6. Next  $h_i$  is sampled in the predicted ellipsoid and checked whether  $w_i = h_i - y_i$  is within its bounds of Eq. 23.  $\rho_i$  and  $c_i$  are determined from the passed samples as shown in Fig 7(a). If the passed samples are divided in some divisions like in Fig 7(b), linearization and sampling are started again for each division and then multiple segmented bounds are obtained. In such cases, multiple updated ellipsoids are generated. Each corresponds to a candidate of interpretation.

## 5 Experimental Result

### 5.1 Experimental Setup

We have performed computer simulation experiments to prove the validity of the method. For simplicity, we use a 2-D link object as in Fig. 8 to estimate its lengths  $r$  and joint angles  $\theta$ . It has a three joints rotating in a  $x - y$  plane and its dynamics is unknown – transition matrix  $\mathbf{A}$  is assumed to be identity. However, all joint angles and the differences between the most proximal joint angle and the second proximal one are constrained within a certain range. Possible ranges of the link lengths are also assumed. As an observation system, only 1-D position  $x$  for each joint is observed and the depth  $y$  is not available. The transition and observation noise is supposed to be bounded.





**Fig. 8.** Representation of an articulated object

**Table 1.** Object constraints (tight set)

pose constraints	$ \theta_2 - \theta_3  \leq \pi/6\text{rad},$ $0\text{rad} \leq \theta_2, \theta_3 \leq \pi/2\text{rad},$
shape constraints	$0 \leq r_1 - r_2 \leq 15,$ $ r_2 - r_3  \leq 15,$ $65 \leq r_1 \leq 90,$ $50 \leq r_2 \leq 75,$ $45 \leq r_3 \leq 65$

**Table 2.** Object constraints (loose set)

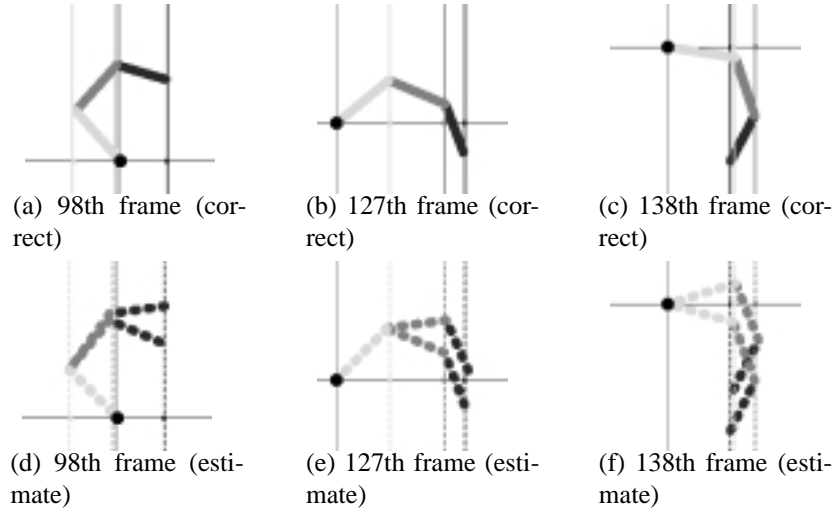
pose constraints	$ \theta_2 - \theta_3  \leq \pi/6\text{rad},$ $0\text{rad} \leq \theta_2, \theta_3 \leq \pi/2\text{rad},$
shape constraints	$0 \leq r_1 - r_2 \leq 30,$ $ r_2 - r_3  \leq 30,$ $52.5 \leq r_1 \leq 102.5,$ $37.5 \leq r_2 \leq 87.5,$ $20 \leq r_3 \leq 80$

## 5.2 Result of Shape Recovery and Accuracy

In order to verify the estimation accuracy, we show estimation results for the link object. First, the tight constraint set shown in Tab. 1 is applied. The result is shown in Fig. 10 and 11. Then another constraint set shown in Tab. 2 twice times looser than Tab. 1 is applied. The result is shown in Fig. 12. With the tight constraints, the estimated range of  $r$  (shown by each vertical bar) is getting small in early time. With the loose constraints in contrast, the range is getting small more slowly but finally converges to the correct value (shown by a straight line). The final range is almost the same as the tight constraints. In Fig. 10, there are cases that more than one estimates (shown as circles) are obtained. This means that there are multiple symmetric interpretations due to depth ambiguity (see Figs. 9(d)-(f)).

## 5.3 Dependency on Initial Estimate

Next, in order to verify dependency of the estimate on initial estimates, we show estimation results for the same observation sequence starting from two different initial estimates in Figs. 11 and 13. Regardless of initial estimates, each shape estimate of  $r_1$  finally converges to almost the same correct value. This means that there is no remarkable dependency on initial estimates.



**Fig. 9.** Multiple estimates

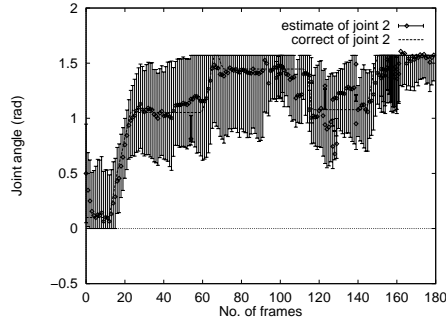
#### 5.4 Identification of Different Shape

In order to verify the ability to identify different shapes, we show estimation results for two objects whose link lengths differ from each other. The estimation for each object starts from the same initial estimate and with the same constraint set. As the experimental result, Fig. 14 shows the correct, initial, and finally estimated objects extending all joints in order to compare lengths of the corresponding links. Concludingly, Fig. 14 tells that all link lengths appropriately converge to the correct values for both objects. This means that the system correctly identified the two different shapes.

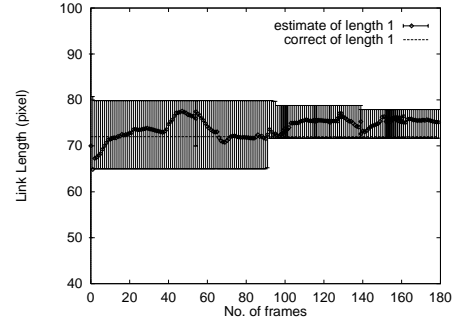
## 6 Conclusion and Discussion

This paper has proposed a method which simultaneously estimates a 3-D shape and moving pose of an articulated object from a monocular image sequence.

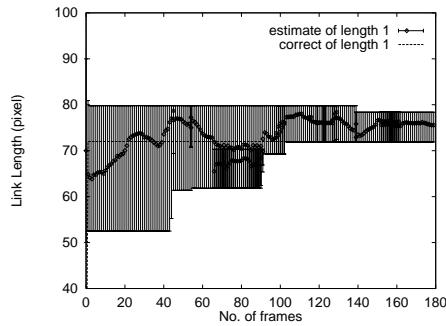
This paper pointed out a Kalman filter estimate essentially converges to wrong state for non-linear unobservable systems. For an alternative, the paper has proposed a set-membership-based method including dynamics. The method limits the depth ambiguity by considering inequality constraints such as possible ranges of joint angles, lengths or widths of parts. Using these constraints, it provides shape recovery of articulated objects under monocular systems. Effectiveness of the framework has been shown by estimation results for an articulated object. Although the system shown in the experiment section adopts an only two-dimensional link object whose joint points can be observed, the proposed method can be applied to cases of three-dimensional objects. The method can use any other observations such as object axes or contours extracted by edge detection or active contour method. A weak point of set-membership methods is an outlier problem. Since bounded noises are supposed, the intersection set may become empty by an outlier noise over the pre-estimated noise bounds. It is a future work to detect and remove outliers.



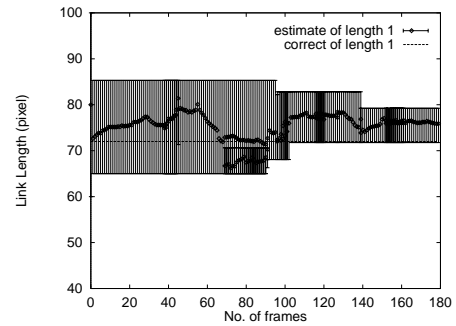
**Fig.10.** Estimate of joint angle  $\theta_2$



**Fig.11.** Estimates of the 1st link length  $r_1$



**Fig.12.** Estimates of the 1st link length  $r_1$  with looser constraints



**Fig.13.** Estimates of the 1st link length  $r_1$  (started from another initial estimate)

## Appendix: Predictive Update by Dynamics

Here, define a vector  $z_t$  as

$$z_t = ((x_t - q_t)^T, u_t^T)^T. \quad (24)$$

With the above equation and Eq. 13, the following relation

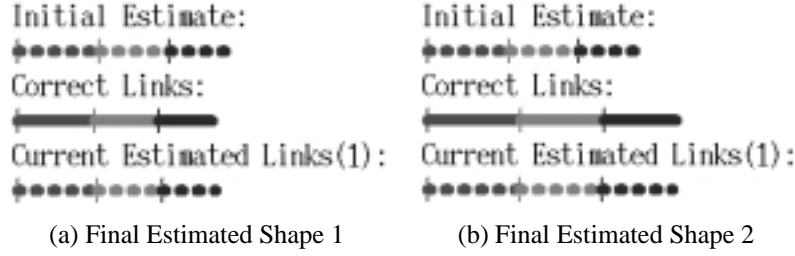
$$z_t^T Q z_t \leq 1 \quad \text{where} \quad Q = \begin{pmatrix} (1 + \lambda)P_t & O \\ O & (1 + \frac{1}{\lambda})U \end{pmatrix}^{-1} \quad (25)$$

is derived and Eqs. 1 and 24 conclude

$$x_{t+1} = A(x_t - q_t) + B u_t + A q_t = (A \ B) z_t + A q_t. \quad (26)$$

In general, the following relations

$$z^T Q^{-1} z \leq 1, \quad x = C z + d \quad (27)$$



**Fig. 14.** Object identification

yield

$$(\mathbf{x} - \mathbf{d})^T (\mathbf{C}\mathbf{Q}\mathbf{C}^T)^{-1} (\mathbf{x} - \mathbf{d}) \leq 1 \quad (28)$$

if  $m \leq n$  and rank  $\mathbf{C}$  equals to  $m$  where  $\mathbf{C}$  is a  $m$  times  $n$  matrix (this condition means that  $\{\mathbf{C}\mathbf{z}\}$  spans a linear space of  $m$  dimensions). Therefore Eq. 13 is rewritten as the form of Eq. 5 and the followings are derived by substituting  $\mathbf{C} = (\mathbf{A} \ \mathbf{B})$ ,  $\mathbf{d} = \mathbf{A}\mathbf{q}_t$  and  $\mathbf{x} = \mathbf{x}_{t+1}$  into Eq. 28:

$$(\mathbf{x}_{t+1} - \mathbf{A}\mathbf{q}_t)^T \left\{ (1 + \lambda)\mathbf{A}\mathbf{P}_t\mathbf{A}^T + \left(1 + \frac{1}{\lambda}\right)\mathbf{B}\mathbf{U}\mathbf{B}^T \right\} (\mathbf{x}_{t+1} - \mathbf{A}\mathbf{q}_t) \leq 1. \quad (29)$$

By comparison of Eq. 29 with Eq.5, Eqs. 14 and 15 are derived.

## References

1. D. Lowe. "Fitting Parameterized Three Dimensional Models to Images". *IEEE Trans., Pattern Anal. Machine Intell.*, vol.13, No.5, pp. 441–450, 1991.
2. R. J. Holt, A. N. Netravali, T. Huang, and Richard J. Qian. "Determining Articulated Motion from Perspective Views: A Decomposition Approach". In *Proc. of Workshop on Motion of Non-Rigid and Articulated Objects '94*, pp. 126–137. IEEE, 1994.
3. N. Shimada, Y. Shirai, Y. Kuno, and J. Miura. "Hand Gesture Estimation and Model Refinement using Monocular Camera – Ambiguity Limitation by Inequality Constraints". In *Proc. of 3rd Int. Conf. on Automatic Face and Gesture Recognition*, pp. 268–273, 1998.
4. F. C. Schweppe. "Recursive State Estimation: Unknown but Bounded Errors and System Inputs". *IEEE Trans. on Automatic Control*, vol.AC-13, No.1, pp. 22–28, 1968.
5. E. Fogel and Y. F. Huang. "On the Value of Information in System Identification – Bounded Noise Case". *Automatica*, vol.18, No.2, pp. 229–238, 1982.
6. Y. Hel-Or and M. Werman. "Recognition and Localization of Articulated Objects". In *Proc. of Workshop on Motion of Non-Rigid and Articulated Objects '94*, pp. 116–123. IEEE, 1994.
7. S. P. Engelson G. D. Hager and S. Atiya. "On Comparing Statistical and Set Based Methods in Sensor Data Fusion". *Proc. IEEE Int. Conf. on Robotics and Automation '93*, pp. 352–358, 1993.
8. L. Chisci, A. Grulli, and G. Zappa. "Recursive State Bounding by Parallelotopes". *Automatica*, Vol. 32 No.7, pp. 1049–1055, 1996.
9. R. E. Ellis. "Geometric Uncertainties in Polyhedral Object Recognition". *IEEE Trans. on Robotics and Automation*, vol.7, No.3, pp. 361–371, 1991.