Proceedings of the Fourth Asian Conference on Computer Vision (ACCV2000), Vol.II, pp.1008–1013 (2000 1)

Extraction of Human Limb Regions and Parameter Estimation based on Curl of Optical Flow

Toru TAMAKI[†‡] * Tsuyoshi YAMAMURA[♭] Noboru OHNISHI[†‡]

[†]Dept. of Info. Eng., Nagoya Univ. Nagoya 464–8603 Japan

[♭]Faculty of Info. Sci. Tec., Aichi Prefectural Univ. Aichi 480–1198 Japan

[‡]Bio-Mimetic Control Research Center, RIKEN Nagoya 463–0003 Japan

ABSTRACT

We propose a method for extracting a rotating limb region and estimating its motion parameters, which is based on the curl of optical flow. Regarding the optical flow as a vector field, the proposed method computes the curl of the optical flow to extract a rotating region. The method assumes that one angular velocity varies in a short time period. By using the EM algorithm, the center of rotation along with rotation angles for each frame is estimated, and simultaneously the rotating region is segmented. Since the proposed method doesn't require a simplicity of background, and can discriminate translational movements produced by other moving objects in the background, it is applicable to a practical situation in which conventional methods don't work well. Experimental results with real image sequences demonstrate that the limb extraction can be performed well from noisy optical flow which are obtained from real images.

Keywords: curl, optical flow, pose estimation, motion segmentation, EM algorithm

1. INTRODUCTION

It is important to extract human regions from a movie as a part of human activity recognition system, such as gesture recognition for human interface, motion reconstruction in virtual reality, and surveillance for security system. So far, methods using markers or instruments that are attached to a subject have been studied[1]. Although they can accurately measure motions of the subject, the problems are that they work under a limited or controlled environment in which the system can be installed, and that they may disturb an action of the subject. On the other hand, an image-based recognition system without any attachments has the advantages that the subject can freely move around without restriction of movements, and that a camera setting is relatively simple even if more than one cameras are used. Hence, it is required to extract a human region from a scene for recognizing his/her actions and postures.

Many studies on recognition of human activities use parameterized human body models because the models enable us to reconstruct the actual human posture, while non-model based methods [2, 3] classify human actions into several patterns and recognize what kind of motion occurs. Some models representing a human body are made of jointed links [4, 5], and some are 2D rectangles[6, 7], volumetric 3D cylinders or polyhedra [8, 9, 10, 11, 12]. Using these models, there are the following two types of model based recognition method. One is that a posed model is projected onto the image plane, and a difference between the projected model and a region of a subject extracted by background subtraction or moving object detection, is evaluated [8, 9, 5, 12]. This method can recognize the posture of the subject from parameters of the model whose silhouette is most similar to the extracted region. However, the region extraction assumes that a background is known or at least uniform color to make the subtraction easy, otherwise there is no moving object except the subject. The other type is a human motion tracking using kinematics of the model [4, 6, 10, 11]. Since the tracking uses optical flow as the information of motion, it isn't affected by the complex static background. But it doesn't extract the region of the subject nor determine the posture at the beginning of tracking, then an interactive tool are required to fit the model on the subject manually.

Therefore, it is significant to extract humans on condition that other moving objects exist and the background is complex and unknown. But, for the difficulties against which the studies mentioned above struggle, it is not easy to accomplish the extraction of whole human body immediately. Therefore, as a first step to the extraction under such a real environment, it is reasonable to detect and extract a human arm in a scene because it is a good key/cue [13] to know where the subject is and what he/she acts, especially for recognition of gesture which is determined by the arm movements.

In this paper, we propose a method to extract a region of a rotating human limb represented by a stick model and estimate its motion parameters. Since the proposed method is based on curl of optical flow, the extraction of rotating limb can be achieved against any background and remove other moving objects. We describe the extraction based on the curl of optical flow in section **2** and the estimation of the model parameters in section **3**. We give an algorithm of the segmentation and the estimation using the EM algorithm in section **4**, and provide experimental results of real images in section **5**.

2. LIMB EXTRACTION BASED ON CURL

In this section, we will explain the property of curl for a vector field produced by rotation, and describe examples of curl for simple and general rotation showing that limb segmentation can be achieved with the curl of optical flow.

2.1 Property of curl

Let $\dot{\boldsymbol{p}}(\boldsymbol{x}) = (u, v, w)$ be the velocity of a point $\boldsymbol{x} = (x, y, z)$ in a three-dimensional space, and these

*e-mail : tamaki@ohnishi.nuie.nagoya-u.ac.jp

velocity vectors construct a vector field. The curl[1] of the vector field is given by:

$$\nabla \times \dot{\boldsymbol{p}}(\boldsymbol{x}) \equiv (w_y - v_z,\, u_z - w_x,\, v_x - u_y)^T , \quad (1)$$

where w_y is $\frac{\partial w}{\partial y}$, and so on.

The curl has the following property when the vector field corresponds to rotation as discussed later.

$$\nabla \times \{\dot{\boldsymbol{p}}(\boldsymbol{x} - \boldsymbol{c}) + \boldsymbol{t}\} = \nabla \times \dot{\boldsymbol{p}}(\boldsymbol{x}) , \quad (2)$$

where $\boldsymbol{c}$ is an arbitrary vector and $\boldsymbol{t}$ is a constant vector field not depending on position. The optical flow in a two-dimensional image plane is regarded as a vector field with $z = w = 0$. From a viewpoint of the optical flow produced by a rotating object, the arbitrary vector $\boldsymbol{c}$ is a center of the rotation, and the constant vector field $\boldsymbol{t}$ is caused by a translation velocity of the object. $\boldsymbol{t}$ is also regarded as velocity caused by translation of a camera or moving objects in a background, and these effects are removed by the property of curl.

2.2 Examples of curl

At first, we consider a simple case as shown in Fig.1(a). This figure is an orthographic projection of a stick rotating on a plane parallel to the image plane. Let $\boldsymbol{p}$ be a point on the limb at a distance r $(0 \le r \le r_1)$ from the center of rotation (the origin) at an angle of θ $(= \omega t)$. The velocity $\dot{\boldsymbol{p}}$ at the point $\boldsymbol{p}$ and the curl $\nabla \times \dot{\boldsymbol{p}}$ are represented as

$$\boldsymbol{p} = (x, y, 0)^T = (r\cos\theta, r\sin\theta, 0)^T \quad (3)$$

$$\begin{aligned}\dot{\boldsymbol{p}} &= (-\omega r\sin\theta, \omega r\cos\theta, 0)^T \\ &= (-\omega y, \omega x, 0)^T \quad (4)\end{aligned}$$

$$\nabla \times \dot{\boldsymbol{p}} = (0, 0, 2\omega)^T \quad (5)$$

Seeing Eq.(5), the z component of the curl of the optical flow is 2ω (twice of the angular velocity), and doesn't depend on neither θ nor r, that is, small velocity around the center of the rotation and large velocity on the top of the arm have the same feature value.

In general, the rotation center isn't necessarily the coordinate origin, and we don't need to know a priori the position of rotation center. Let a point $\boldsymbol{c} = (c_x, c_y)^T$ be the center of the rotation of the stick. Then, the velocity $\dot{\boldsymbol{p}} = (u, v)^T$ at a point $\boldsymbol{p} = (x, y)^T$ on the image is represented as follows.

$$\dot{\boldsymbol{p}} = \begin{pmatrix} u \\ v \\ 0 \end{pmatrix} = \begin{pmatrix} -\omega(y - c_y) \\ \omega(x - c_x) \\ 0 \end{pmatrix} \quad (6)$$

$$\nabla \times \dot{\boldsymbol{p}} = (0, 0, v_x - u_y)^T = (0, 0, 2\omega)^T \quad (7)$$

The z component of the curl Eq.(7) is identical with Eq.(5).

In the discussion above, the limb is rotating on a plane parallel to the image plane. As the first extension of the method to deal with 3D limb movements, we consider the case that the limb rotates on a plane not parallel to the image plane.

As shown in Fig.1(b), the limb rotates on the plane which is the one after rotating the image plane (O-xy) by an angle of ϕ around the y axis. The velocity $\dot{\boldsymbol{p}}$ of a point $\boldsymbol{p}$ on the limb is represented as;

[1] Though the term "rotation"or "rot" also means $\nabla \times \dot{\boldsymbol{p}}(\boldsymbol{x})$, we don't use them because it is confusing with a physical rotation of an object around an axis.

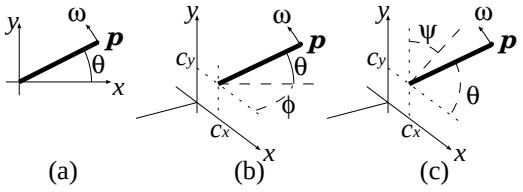

Figure 1: The stick model moving on a plane parallel to the image plane (a), and the plane (O-xy) rotates around (b) the y axis and (c) the x axis.

$$\begin{aligned}\boldsymbol{p} &= (x, y, z)^T \\ &= (r\cos\phi\cos\theta, r\sin\theta, r\sin\phi\cos\theta)^T \quad (8)\end{aligned}$$

$$\begin{aligned}\dot{\boldsymbol{p}} &= (-\omega r\cos\phi\sin\theta, \omega r\cos\theta, -\omega r\sin\phi\sin\theta) \\ &= (-\omega\cos\phi\, y, \frac{\omega x}{\cos\phi}, -\omega\sin\phi\, y),\ \phi \ne \frac{\pi}{2} \quad (9)\end{aligned}$$

When the rotation center is $\boldsymbol{c} = (c_x, c_y, c_z)^T$, the velocity $\dot{\boldsymbol{p}} = (u, v, w)^T$ at a point $\boldsymbol{p} = (x, y, z)^T$ is represented as follows.

$$\dot{\boldsymbol{p}} = \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} -\omega\cos\phi(y - c_y) \\ \dfrac{\omega(x - c_x)}{\cos\phi} \\ -\omega\sin\phi(y - c_y) \end{pmatrix} \quad (10)$$

The z component of the curl is;

$$v_x - u_y = \omega\frac{1 + \cos^2\phi}{\cos\phi} \quad (11)$$

Assuming that ϕ is constant in a short time period (like frame rate) even if ϕ varies in the actual limb movement, we consider that Eq.(11) also depends on only ω.

When the limb rotates on the plane at an angle of ψ around the X axis shown in Fig.1(c), a similar result is obtained as follows.

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} \dfrac{-\omega(y - c_y)}{\cos\psi} \\ \omega\cos\psi(x - c_x) \\ \omega\sin\psi(x - c_x) \end{pmatrix} \quad (12)$$

$$v_x - u_y = \omega\frac{1 + \cos^2\psi}{\cos\psi} \quad (13)$$

In general, the angles of the plane on which the limb rotates are $\phi \ne 0$ and $\psi \ne 0$. The velocity becomes a little complicated, but the curl is still simple.

$$v_x - u_y = \omega\frac{\cos^2\psi + \cos^2\psi + \sin\phi\sin\psi}{\cos\phi\cos\psi} \quad (14)$$

In the following discussions, we will focus mainly on the case of ϕ.

2.3 Segmentation with curl

As mentioned above, the z components of curl shown in Eqs.(11), (13) and (14) depend on ω on the condition that ϕ and ψ are constant for an instant. Under orthographic projection, (u, v) of the velocity $\dot{\boldsymbol{p}}$ is regarded as optical flow, and Eqs.(11), (13) and (14) can be calculated from optical flow. Therefore, using the z component of curl, we can segment region corresponding to rotating limb with the angular velocity ω. Even if several objects rotate in a scene, the regions of the objects can be segmented as long as the angular velocities of the objects are different from one another.

2.4 Under perspective projection

The discussions above assume the orthographic projection because of the simplification, and it is reasonable when the motion along the optical axis of a camera is small as compared with the distance between the camera and the object. Note that, however, the z component of curl under a perspective projection is the same with that of the orthogonal projection as long as the rotation plane is parallel to the image plane.

3. ESTIMATING MOTION PARAMETERS

In the previous section, we mentioned the method to extract the region of the rotating limb on the slant plane based on the curl of optical flow. In this section we discuss how to estimate the motion parameters; the center of rotation (c_x, c_y), the angular velocity ω, the slant angle ϕ, and the limb angle θ.

Since optical flow computed from real images involves inevitable noise, we should use the least square method to estimate the parameters from the optical flow within an extracted region. We make simultaneous equations from Eq.(10) in which the velocity $\dot{\boldsymbol{p}}_j=(u_j, v_j)^T$ of a point $\boldsymbol{p}_j=(x_j, y_j)^T$ in the region is represented by the parameters ω, ϕ, c_x, c_y of the region.

However, it is not easy to solve the simultaneous equations directly because of the non-linearity. Hence, we transform Eq.(10) as follows.

$$\begin{pmatrix} u_j \\ v_j \end{pmatrix} = \begin{pmatrix} \alpha y_j + \beta \\ \gamma x_j + \delta \end{pmatrix} \tag{15}$$

$$\alpha=-\omega\cos\phi,\ \beta=\omega c_y\cos\phi,\ \gamma=\frac{\omega}{\cos\phi},\ \delta=\frac{-\omega c_x}{\cos\phi} \tag{16}$$

This is rewritten as;

$$\dot{\boldsymbol{p}}_j = A_j\boldsymbol{q} \tag{17}$$

$$A_j = \begin{pmatrix} y_j & 1 & 0 & 0 \\ 0 & 0 & x_j & 1 \end{pmatrix}, \quad \boldsymbol{q} = (\alpha, \beta, \gamma, \delta)^T \tag{18}$$

Eq.(17) shows that the optical flow is represented by the linear functions of the parameters α, β, γ, and δ. Therefore, we construct the following simultaneous linear equations.

$$\begin{pmatrix} \dot{\boldsymbol{p}}_1 \\ \dot{\boldsymbol{p}}_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \end{pmatrix} \boldsymbol{q} \tag{19}$$

We solve this over-determined equations with singular value decomposition (SVD) [14] as a least square problem. SVD divides a $m\times n$ $(m>n)$ matrix A into UWV^T, where U is a $m\times n$ column orthogonal matrix, W is a $n\times n$ diagonal matrix whose elements are positive, and V is a $n\times n$ orthogonal matrix. Representing Eq.(19) as $\dot{\boldsymbol{p}}=A\boldsymbol{q}$, the solution is obtained as $\boldsymbol{q}=VW^{-1}U^T\dot{\boldsymbol{p}}$, and the parameters are calculated from $\alpha, \beta, \gamma, \delta$ as follows.

$$c_x=\frac{-\delta}{\gamma},\ c_y=\frac{-\beta}{\alpha},\ |\omega|=\sqrt{-\alpha\gamma},\ \cos\phi=\sqrt{\frac{-\alpha}{\gamma}} \tag{20}$$

For ω, the rotation movement should subject to $-\alpha\gamma\geq 0$. Of course, $0\leq\frac{-\alpha}{\gamma}\leq 1$ (first for the square root and second for cosine) also should be satisfied for ϕ. We will discuss about these constraints later along with the sign of ω.

This method can estimate the parameters except the angle of the limb θ because Eq.(19) doesn't have any information about θ. Therefore, we estimate θ by finding the principal axis of the extracted limb, but this must be replaced by a more sophisticated way; model fitting and the like.

4. SEGMENTATION WITH THE EM ALGORITHM

In section 3, we have presented the method to estimate the parameters of the limb using optical flow of the region which is extracted in advance by the segmentation with the z component of curl of optical flow.

The estimation of the motion parameters can be done when a region is determined, while the limb segmentation requires an accurate computation of the curl. However, it is difficult to compute the curl accurately at each point in real images because curl is a differential operation and sensitive to noise. Seeing Eq.(15), we find that optical flow can be estimated by using the obtained parameters α, β, γ, and δ. Then we use the gradients of the estimated optical flow instead of the values calculated by the differential operation.

The problem here is that the estimation and the segmentation depend on each other; the parameter estimation needs accurate segmentation, while the region segmentation requires accurate parameter values. To overcome this problem, we use the expectation and maximization (EM) algorithm [15], which is applicable to the maximum likelihood estimation of mixture distributions involving unobserved variables. When several distributions compose a mixture distribution, we don't know a distribution to which each data belongs. The unobserved variable determines a distribution and makes data belong to it, and the segmentation of the data is performed.

4.1 The algorithm for segmentation

We assume that the distribution of the error between $\dot{\boldsymbol{p}}_j$ and $A_j\boldsymbol{q}$ is subject to a two-dimensional Gaussian. Hence, we define the following conditional probability of $\dot{\boldsymbol{p}}_j$ at $\boldsymbol{p}_j$ with $\boldsymbol{q}=(\alpha, \beta, \gamma, \delta)^T$.

$$P(\dot{\boldsymbol{p}}_j|\ \boldsymbol{p}_j, \boldsymbol{q}, \Sigma) = \frac{1}{2\pi|\Sigma|^{\frac{1}{2}}}\exp\left\{\frac{-1}{2}\Big(\dot{\boldsymbol{p}}_j - A\boldsymbol{q}\Big)^T\Sigma^{-1}\Big(\dot{\boldsymbol{p}}_j - A\boldsymbol{q}\Big)\right\} \tag{21}$$

where $\Sigma=\begin{pmatrix}\sigma_x^2 & 0 \\ 0 & \sigma_y^2\end{pmatrix}$ is a covariance matrix. Here we assume that the errors for u $(=\alpha y+\beta)$ and v $(=\gamma x+\delta)$ are independent of each other because α, β and γ, δ are estimated separately, and Eq.(21) is transformed into the product of two probabilities.

$$P(\boldsymbol{p}_j|\ \dot{\boldsymbol{p}}_j, \boldsymbol{q}, \sigma_x^2, \sigma_y^2) = \frac{1}{\sqrt{2\pi\sigma_x^2}}\exp\left\{\frac{-(u_j-\alpha y_j-\beta)^2}{2\sigma_x^2}\right\} \cdot \frac{1}{\sqrt{2\pi\sigma_y^2}}\exp\left\{\frac{-(v_j-\gamma x_j-\delta)^2}{2\sigma_y^2}\right\} \tag{22}$$

We show below the algorithm to segment the limb region and estimate its parameters using the probability defined by Eq.(22).

1. Compute optical flow $\dot{\boldsymbol{p}}_j = (u_j, v_j)^T$ at each point $\boldsymbol{p}_j = (x_j, y_j)^T$ $(j = 1, \ldots, N)$.
2. Classify the optical flow based on the direction of velocity to obtain cluster R_i $(i = 1, \ldots, M)$.

Table 1: The estimation result at different ϕ ($\omega = 1$).

ϕ ($\cos\phi$)	0 (1)	10 (0.9848)	20 (0.9397)	30 (0.8660)	[deg]
estimates of ϕ ($\cos\phi$)	7.18 (0.9923)	7.97 (0.9903)	15.8 (0.9624)	25.2 (0.9048)	[deg]
estimates of ω	0.913	0.939	0.936	0.948	[deg/frame]

3. Let weight w_{ij} represent a probability that a point $\boldsymbol{p}_j$ belongs to a cluster R_i with parameters of $\boldsymbol{q}_i = (\alpha_i, \beta_i, \gamma_i, \delta_i)$. Set the initial values of w_{ij} as follows.

$$w_{ij} = \begin{cases} 1 & (\boldsymbol{p}_j \in R_i) \\ 0 & (\boldsymbol{p}_j \notin R_i) \end{cases} \tag{23}$$

4. Normalize w_{ij} so that the sum of them for the cluster R_i is 1.

$$w_{ij} \leftarrow \frac{w_{ij}}{\xi_i}, \quad \text{where} \quad \xi_i = \frac{1}{N}\sum_j w_{ij} \tag{24}$$

5. Find the parameters $\boldsymbol{q}_i = (\alpha_i, \beta_i, \gamma_i, \delta_i)$ of each cluster R_i minimizing the following weighted square errors : $\sum_j w_{ij}||\dot{\boldsymbol{p}}_j - A_j\boldsymbol{q}_i||^2 = \sum_j ||(\sqrt{w_{ij}}\dot{\boldsymbol{p}}_j) - (\sqrt{w_{ij}}A_j)\boldsymbol{q}_i||^2$

   We solve the following simultaneous equations by SVD to obtain the parameters $\boldsymbol{q}_i$.

$$\begin{pmatrix} \sqrt{w_{i1}}\,\dot{\boldsymbol{p}}_1 \\ \sqrt{w_{i2}}\,\dot{\boldsymbol{p}}_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} \sqrt{w_{i1}}A_1 \\ \sqrt{w_{i2}}A_2 \\ \vdots \end{pmatrix} \boldsymbol{q}_i \tag{25}$$

6. Compute the weighted variances $\sigma_{xi}^2, \sigma_{y_i}^2$ of each cluster R_i.

$$\sigma_{xi}^2 = \frac{1}{N}\sum_j w_{ij}(u_j - \alpha_i y - \beta_i)^2 \tag{26}$$

$$\sigma_{y_i}^2 = \frac{1}{N}\sum_j w_{ij}(v_j - \gamma_i y - \delta_i)^2 \tag{27}$$

7. Update w_{ij} with the equation (28) by using $\boldsymbol{q}_i$, σ_{xi}^2, and $\sigma_{y_i}^2$ obtained from Eqs.(25)$\sim$(27);

$$w_{ij} = \frac{\xi_i P(\dot{\boldsymbol{p}}_j|\ \boldsymbol{p}_j, \boldsymbol{q}_i, \sigma_{xi}^2, \sigma_{y_i}^2)}{\sum_k \xi_k P(\dot{\boldsymbol{p}}_j|\ \boldsymbol{p}_j, \boldsymbol{q}_k, \sigma_{xk}^2, \sigma_{y_k}^2)} \tag{28}$$

8. If $\sum_i ||\boldsymbol{q}_i^{new} - \boldsymbol{q}_i^{old}||^2 > \varepsilon$, that is, the difference between the estimated parameter values of this iteration $\boldsymbol{q}_i^{new}$ and that of the previous iteration $\boldsymbol{q}_i^{old}$ is larger than a threshold ε, go to step 4. Otherwise, go to step 9.
9. Obtain region segmentation by making each point $\boldsymbol{p}_j$ belong to the cluster R_i with the largest weight w_{ij}.

4.2 Removing non-rotational movement

By the segmentation mentioned in 4.1, the parameters $\boldsymbol{q}$ are obtained, and c_x, c_y, ω and ϕ are calculated by Eq.(20). For rotation movement, $-\alpha\gamma \geq 0$ should be satisfied so that ω has a real value solution. However, α and γ are estimated separately without checking whether the movement is rotation or not, and as a result, there may exist some regions not satisfying the constraint. Such regions correspond to non-rotational movement; translation and expansion / contraction due to zoom in/out.

Table 2: Motion classification with the signs of the parameters.

constraint	*gradients*		*corresponding motion*
$-\alpha\gamma > 0$	$\alpha > 0$	$\gamma < 0$	rotation ($\omega < 0$)
	$\alpha < 0$	$\gamma > 0$	rotation ($\omega > 0$)
$-\alpha\gamma = 0$	$\alpha = 0$	$\gamma = 0$	translation
$-\alpha\gamma \leq 0$	$\alpha \geq 0$	$\gamma \geq 0$	expansion
	$\alpha \leq 0$	$\gamma \leq 0$	contraction

Table 3: which axis does the rotation plane slant.

around	*criteria*	*parameters*
Y axis	$\lvert\alpha\rvert \leq \lvert\gamma\rvert$	$\cos\phi = \sqrt{\frac{-\alpha}{\gamma}}$, $\psi = 0$
X axis	$\lvert\alpha\rvert \geq \lvert\gamma\rvert$	$\phi = 0$, $\cos\psi = \sqrt{\frac{-\gamma}{\alpha}}$

Table 2 shows the classification of motion based on the sign of $-\alpha\gamma$ (i.e. α and γ). When α and γ have different signs, i.e. $-\alpha\gamma$ is larger than 0, then the corresponding motion is rotation, and the direction of rotation is determined by the sign of α ($= -\omega\cos\phi$) because $\cos\phi \geq 0$. If both of α and γ are 0, the optical flow of the region is constant. This means the region translates or doesn't move. Finally we consider the case of $-\alpha\gamma \leq 0$. When α and γ are positive, the motion is expansion, and when both are negative, the motion is contraction. Thus based on the sign of α and γ, we can select the regions of rotating object removing regions with non-rotational movement.

4.3 Determination of rotation plane (rotation axis)

We can select regions corresponding to rotating objects by using the method described above. Here, we explain how to determine the rotation plane (rotation axis) : the angle between the rotation plane and the image plane. In 2.2, we have investigated two kinds of general rotation shown in Fig.1 and found the following relations among the angle and the parameters α and γ : $\cos\phi = \sqrt{-\alpha/\gamma}$ for the rotation in Fig.1(b), and $\cos\psi = \sqrt{-\gamma/\alpha}$ for the rotation in Fig.1(c) (derived from Eq.(12) with the same manner as Eq.(10)).

Then by using these relations, we can easily determine the angle between the rotation plane and the image plane, i.e., the rotation axis. Table 3 shows how to determine the angle. If $|\alpha| \leq |\gamma|$, the rotation is on the plane which makes the angle ϕ with the image plane, and the rotation axis is $(\sin\phi, 0, \cos\phi) = \left(\sqrt{1-(-\alpha/\gamma)}, 0, \sqrt{-\alpha/\gamma}\right)$. Otherwise the rotation axis is $(0, \sin\psi, \cos\psi)$.

5. EXPERIMENTS AND RESULTS

The proposed method has been implemented on a UNIX workstation with C++. The computation of optical flow (step 1. of the algorithm) is performed by the codes available from [16, 17], and the

initial clusters (step 2.) are made by a simple histogram clustering which divides the directions of velocity vector into 24 sections and finds peaks in the direction histogram as cluster centers.

5.1 Evaluating Estimation Accuracy using Manipulator

At first to evaluate the estimation accuracy of the proposed method, we used a manipulator whose angles can be controlled precisely. For each fixed ϕ of 0, 10, 20 and 30[deg], the images are taken by rotating the link at $\omega = 1$[deg/frame].

The estimates are in Table 1. The angular velocity $\omega(=1)$ is estimated about 0.9∼0.95 and the error is about 10%, and this could be improved by considering results of many frames. On the other hand, the estimate of ϕ has a relatively large error (though $\cos\phi$ is good), so the method for estimating ϕ should be improved.

5.2 Extraction of human limb

We show the results for a human arm movement. The motion takes place at a distance (about 3m) from the camera, and images are taken 10 frames per second, then the size is reduced to 1/4 in order to eliminate the effect of interlace.

Figure 2 shows the original image at 33th frame, the optical flow and the time development of segmentation. At the initial clustering, the arm is divided into the upper and lower regions, but after several iterations the whole arm is covered with one region from the wrist to the shoulder. The drastic change of the segmentation occurs only in early stage of the iterations, and after 10 iterations the segmentation results don't change so much (Fig.2(g) and (h)).

Supposing that an arm region has a largest angle velocity in the image, we extract the regions (see Fig.3) from a sequence of 30th∼38th frames.

Although the estimated parameters involve the error, we can find a tendency of the arm motion through a change of the parameters. The trajectories of the estimated parameters from 30th through 39th frame are shown in Fig.4. We can find that ω becomes smaller as the arm swings down, while ψ doesn't change abruptly.

The other results are shown in Fig.5. Two arms move simultaneously in Fig.5(a), and the two regions of the arms are separately extracted in Fig.5(b) in which the regions satisfy the constraint in Table 2 and have angular velocities larger than a threshold (about 3 deg/s). Fig.5(c) shows that the arm moves in front of the man walking left. Both are moving, but the man which is not rotation is rejected by the constraint and only the arm region remains.

6. CONCLUSIONS

We have proposed the method to extract the rotating limb region based on the curl of optical flow and estimate the parameters of the limb. The proposed method is composed of the segmentation based on the gradients of optical flow, the estimation of the motion parameters, and the EM algorithm which performs the estimation and the segmentation simultaneously. The method separate rotational movements from translation movements in a scene based on the estimated parameters. The experimental results demonstrate that the proposed method can extract clearly the region of the rotating arm from noisy optical flow in real images, though the estimation contains some errors. We will examine the method under conditions that a camera moves and other moving objects exist.

Lastly we briefly describe some extensions of the method. First, a general rotation movement of the limb ($\phi\neq0$, $\psi\neq0$) can be linearized by approximating both u and v with general planes $\alpha x+\beta y+\gamma$ instead of Eq.(15). Secondly, the effect of the constant vector field $\boldsymbol{t}$ in Eq.(2) can be compensated by two set of parameters obtained in consecutive frames, on the assumption that $\boldsymbol{t}$ doesn't vary in a short time period. Finally, when the arm is modeled by two links with an elbow joint, the angular velocity of the parent link which is added to that of its child link will be compensated by regarding the angular velocity as a translation of the child link. These extensions listed above are the future works.

REFERENCES

[1] L.Goncalves, E.D.Bernardo, P.Perona : "Reach Out and Touch Space (Motion Learning)," *Proc. of FG'98*, pp.234–239 (1998).

[2] J.Yamato, J.Ohta, K.Ishii : "Recognizing Human Action in Time-Sequential Images using Hidden Markov Model," *Proc. of CVPR'92*, pp.379–385 (1992).

[3] T.Nishimura, T.Mukai, R.Oka : "Non-monotonic Continuous Dynamic Programming for Spotting Recognition of Hesitated Gestures from Time-Varying Images," *Proc. of ACCV'98*, Vol.2, pp.734–739 (1998).

[4] D.D.Morris, J.M.Rehg : "Singularity Analysis for Articulated Object Tracking," *Proc. of CVPR'98*, pp.289–296 (1998).

[5] Y.Guo, G.Xu, S.Tsuji : "Understanding Human Motion Patterns," *Proc. of ICPR'94*, pp.325–329 (1994).

[6] S.X.Ju, M.J.Black, Y.Yacoob : "Cardboard People: A Parameterized Model of Articulated Image Motion," *Proc. of FG'96*, pp.38–44 (1996).

[7] I.Haritaoglu, D.Harwood, L.S.Davis : "W^4: Who? When? Where? What? A Real Time System for Detecting and Tracking People," *Proc. of FG'98*, pp.222–227 (1998).

[8] D.M.Gavrila, L.S.Davis : "3-D model-based tracking of humans in action: a multi-view approach," *Proc. of ICPR'96*, pp.73-80 (1996).

[9] J.Ohya, F.Kishino : "Human Posture Estimation from Multiple Images Using Genetic Algorithm," *Proc. of ICPR'94*, pp.750–753 (1994).

[10] M. Yamamoto, K. Koshikawa: "Human motion analysis based on a robot arm model," *Proc. of CVPR'91*, pp.664–665 (1991)

[11] M.Yamamoto, A.Sato, S.Kawada, T.Kondo, Y. Osaki : "Incremental tracking of human actions from multiple views," *Proc. of CVPR'98*, pp.2–7 (1998).

[12] M.Minoh : "3D Model Centered Framework for CV and VR," *Proc. of ACCV'98*, Vol.2, pp.332–339 (1998)

[13] L.Goncalves, E.D.Bernardo, E.Ursella, P.Perona : "Monocular tracking of the human arm in 3D," *Proc. of ICCV'95*, pp.764–770 (1995).

[14] W.H.Press, S.A.Teukolsky, W.T.Vetterling, B.P. Flannery : *Numerical recipes in C*, Cambridge University Press (1992).

[15] A.P.Dempster, N.M.Laird, D.B.Rubin : "Maximum Likelihood from Incomplete Data via the EM Algorithm," *J. Roy. Statist. Soc. B*, Vol.39, pp.1–38 (1997).

[16] N.Ohta : `http://www.ail.cs.gunma-u.ac.jp/Labo/Program/Flow-R00.tar.gz`

[17] N.Ohta : "Image movement detection with reliability indices," *IEICE Transactions*, Vol.E74, No.10, pp.3379–3388 (1991).

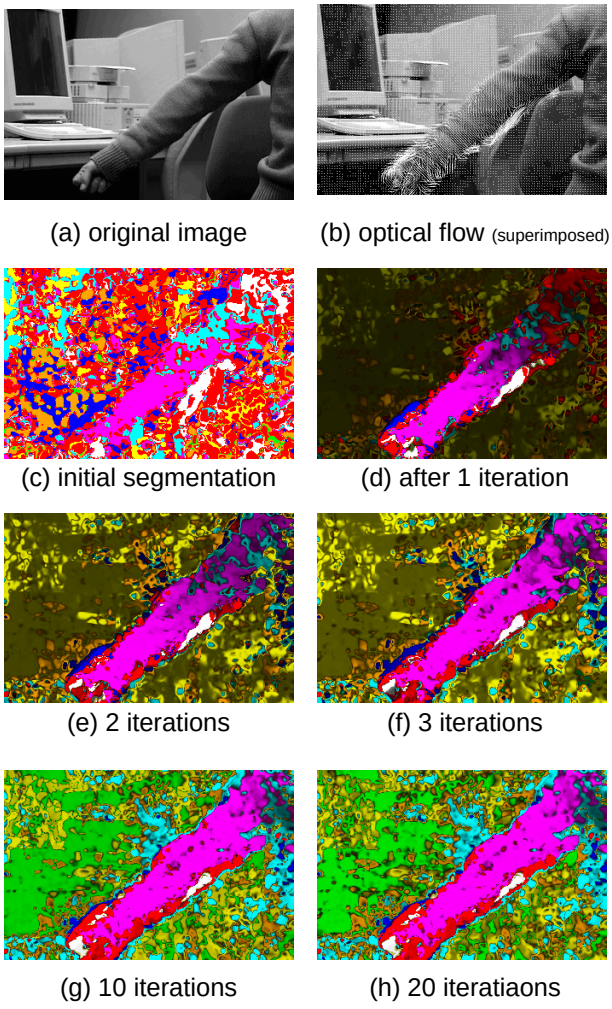

Figure 2: The experimental results using real images (swing downward from horizontal). **(a)** The original image at 33th frame (340×223). **(b)** The optical flow. **(c)**$\sim$**(h)** The segmentation results painted with a different color for each cluster, and the intensity becomes brighter in proportion to the weight w_{ij}.

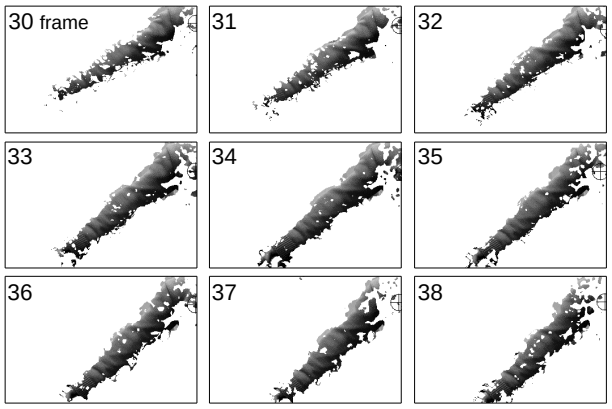

Figure 3: The result of the extraction of the arm over the sequence from 30th frame to 38th. The mark of $\oplus$ indicates a center of rotation.

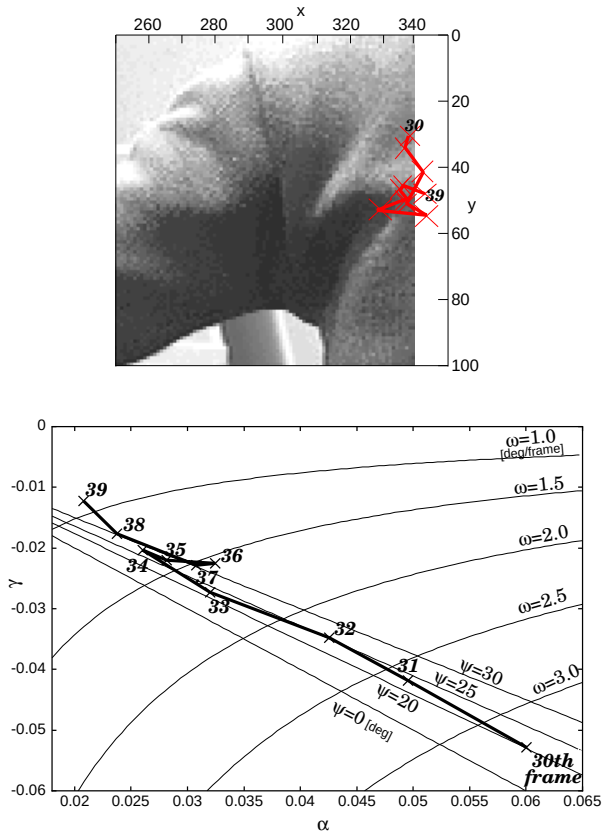

Figure 4: The Trajectory of the estimated values ω and ψ via α and γ. This shows how the estimates change over the frames. **(top)** The trajectory of rotation center is superimposed on the image of 30th frame magnified around the shoulder. The center comes down, and stays in the latter frames. **(bottom)** ω becomes smaller as the arm swings down, while ψ doesn't change abruptly. The thin lines represent curves of $\alpha\gamma = -\omega^2$ and lines of $\alpha = -\cos^2\phi\gamma$.

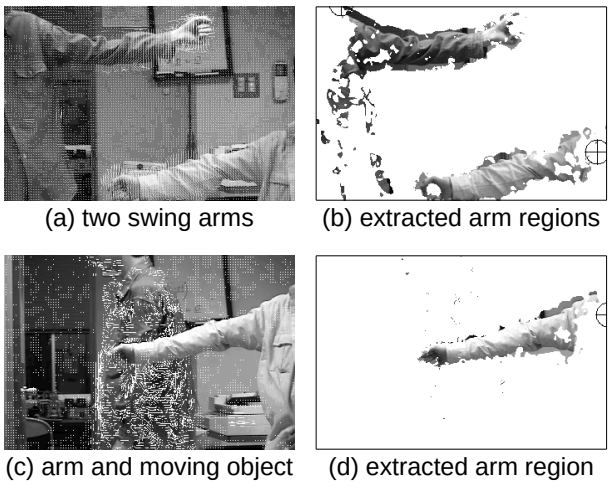

Figure 5: **(a)(b)** Other result where there are two moving arms. **(c)(d)** Other result where moving object exists behind moving arm.