Extracting Human Limb Region using Optical Flow and Nonlinear Optimization

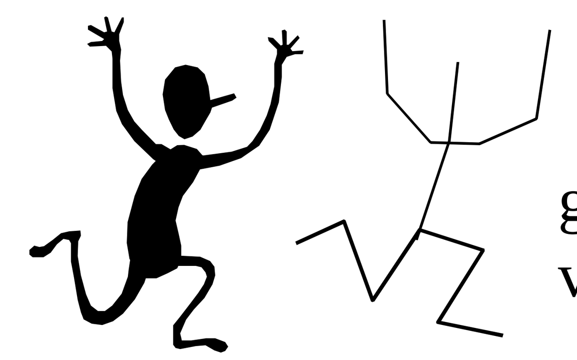
Toru Tamaki (*Niigata Univ.*)
Tsuyoshi Yamamura (*Aichi Pref. Univ.*)
Noboru Ohnishi (*Nagoya Univ. & RIKEN*)

BACKGROUND

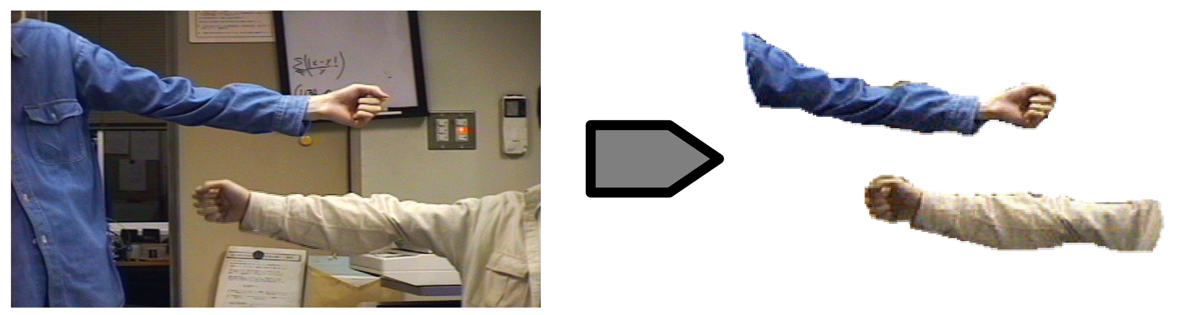
Condition : ***complex background***
multiple moving objects

Extracting human regions from a sequence of images

Extracting human limbs
Estimating motion parameters

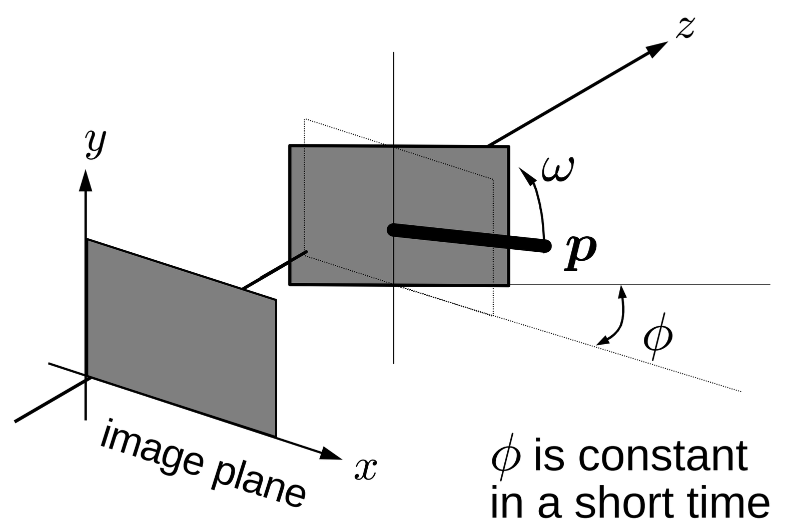


Applications :
gesture recognition
virtual reality



INDIRECT METHOD

Segmentation of optical flow modeled by rotation model with four parameters $\boldsymbol{q}$



ϕ is constant in a short time

Rotation model

$\dot{\boldsymbol{p}}$: optical flow

$\boldsymbol{q}$: the parameters

$$\dot{\boldsymbol{p}}_j = \begin{pmatrix} u_j \\ v_j \end{pmatrix} = \begin{pmatrix} -w(y_j - c_y)\cos\phi \\ \dfrac{w(x_j - c_x)}{\cos\phi} \end{pmatrix} \equiv \begin{pmatrix} a\,y_j + b \\ g\,x_j + d \end{pmatrix} = \begin{pmatrix} y_j & 1 & 0 & 0 \\ 0 & 0 & x_j & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ g \\ d \end{pmatrix} = A_j \boldsymbol{q}$$

Least square solution

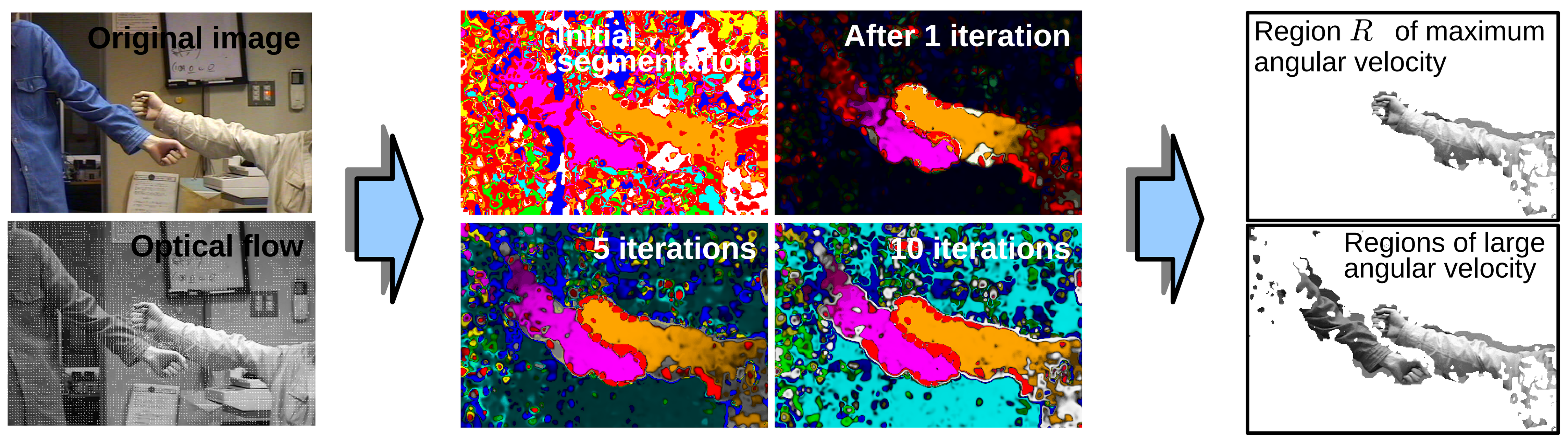
$$\begin{pmatrix} \dot{\boldsymbol{p}}_1 \\ \dot{\boldsymbol{p}}_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \end{pmatrix} \boldsymbol{q} \quad \Rightarrow \quad c_x = \frac{-d}{g}, \quad c_y = \frac{-b}{a}, \quad \omega = -\mathrm{sign}(a)\sqrt{-ag} \quad (-ag \geq 0)$$

(c_x, c_y) : rotation center
ω: angular velocity

Segmentation using iterative method

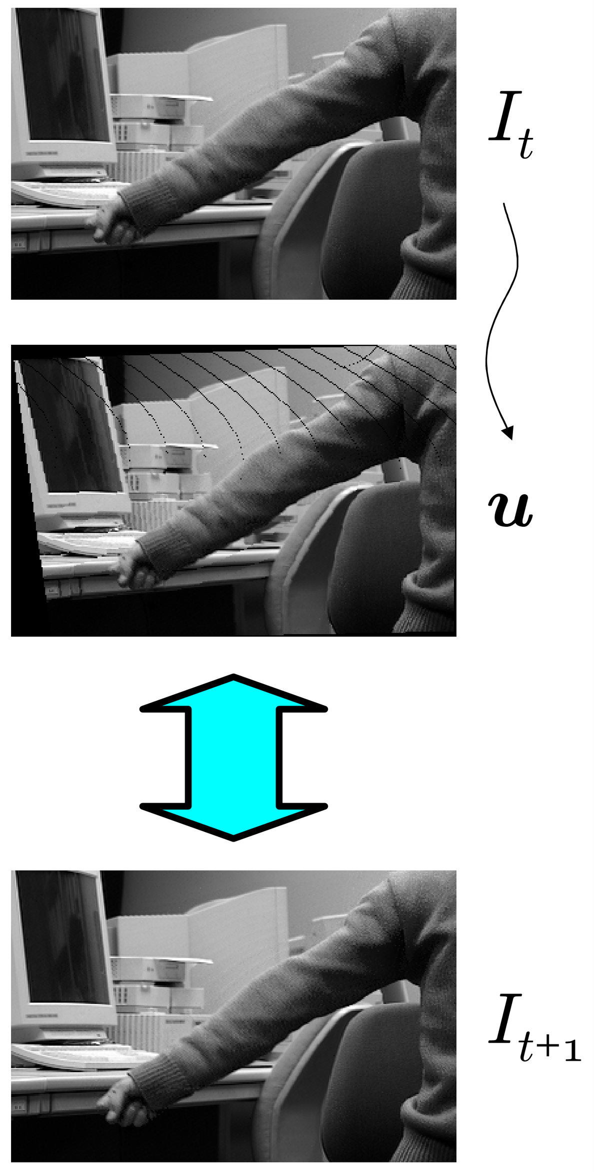
Divide into several regions to apply least square for each moving object $\boldsymbol{q}_i$ by using the EM algorithm

Extraction of Two Arms



DIRECT METHOD

Estimating the motion parameters by minimizing the residuals between two successive images



Residual of intensities

$\boldsymbol{p}_j$: point on the arm, $I_t(\boldsymbol{p})$: intensity at $\boldsymbol{p}$

$$r_j = I_t(\boldsymbol{p}_j) - I_{t+1}(\boldsymbol{p}_j + \boldsymbol{u}(\boldsymbol{p}_j, \boldsymbol{\theta}))$$

Motion model with eight parameters

$\boldsymbol{u}$: motion vector, $\boldsymbol{\theta}$: motion parameters

$$\boldsymbol{u}(\boldsymbol{p}(x,y);\boldsymbol{\theta}) = \begin{pmatrix} 1 & x & y & 0 & 0 & 0 & x^2 & xy \\ 0 & 0 & 0 & 1 & x & y & xy & y^2 \end{pmatrix} \boldsymbol{\theta} = M\boldsymbol{\theta}$$

Estimation by nonlinear optimization

Cost function : $\min_{\boldsymbol{\theta}} \sum_{\boldsymbol{p}_j \in R} r_j^2$ $\quad R$: arm region

Update parameters

$\boldsymbol{\theta} \leftarrow \boldsymbol{\theta} + d\boldsymbol{\theta}$: Gauss-Newton Method

$d\boldsymbol{\theta}$: Decent direction is the solution of

$$\sum_l \sum_j \frac{\partial r_j}{\partial q_k} \frac{\partial r_j}{\partial q_l} dq_l = -\sum_j \frac{\partial r_j}{\partial q_k}$$

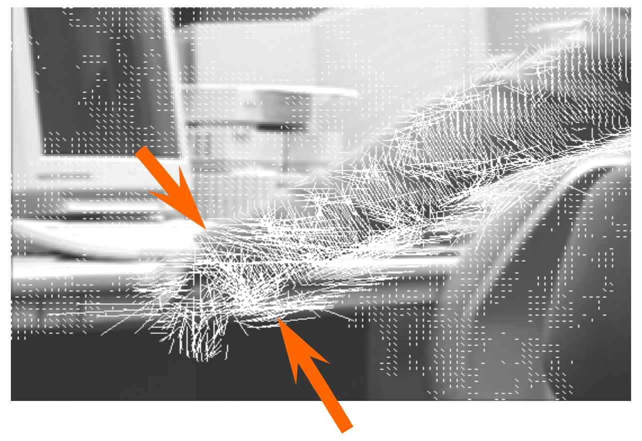
where $\dfrac{\partial r}{\partial q} = \dfrac{\partial \boldsymbol{u}}{\partial q} \dfrac{\partial r}{\partial \boldsymbol{u}} = -M^T \nabla I_{t+1}(\boldsymbol{p} + \boldsymbol{u}(\boldsymbol{p}))$

PROBLEMS

▸ Indirect method

Difficult to compute flow accurately because of the discontinuity of motion

Optical Flow of real image



▸ Direct method

High computational cost because of nonlinear iterative method

COMBINED USE

Image sequence → **Optical flow** → **Indirect method**

Indirect method
- ○faster, robust to flow noise
- ×can't deal with motion edge

Direct method
- ○accurate extraction
- ×high computational cost

- ▸ Extracted region R
- ▸ Estimated parameters q are used as initial state

Computational cost is reduced.
The initial states are given.
Motion edge is allowed.

EXTRACTION ALGORITHM

Probability for each residual r_j

$$P(r_j\,|\,\theta_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{r_j^2}{2\sigma^2}\right)$$

Prior probabilities

Arm region θ_1: $P(\theta_1) = \frac{|R|}{N}$ (area of arm region) (all pixels)

Background θ_0: $P(\theta_0) = 1 - P(\theta_1)$

Decision by MAP

$$P(\theta_1)P(r_j\,|\,\theta_1) > P(\theta_0)P(r_j\,|\,\theta_0)$$

→ Pixel p_j belongs to the arm region

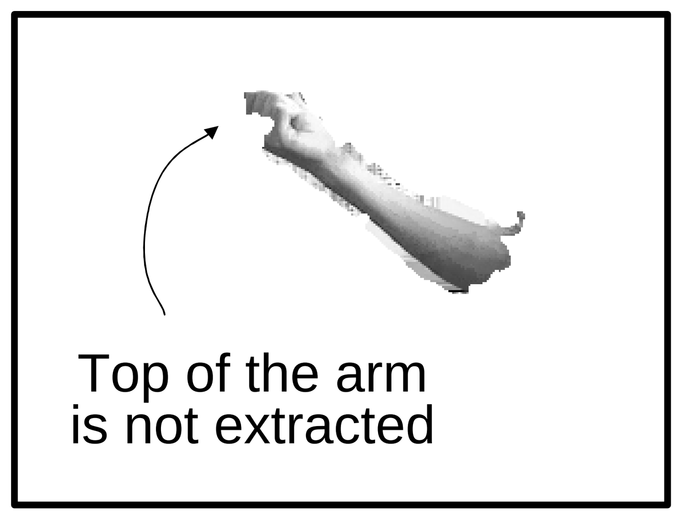
Post processing

Burying holes and smoothing contour of the arm region

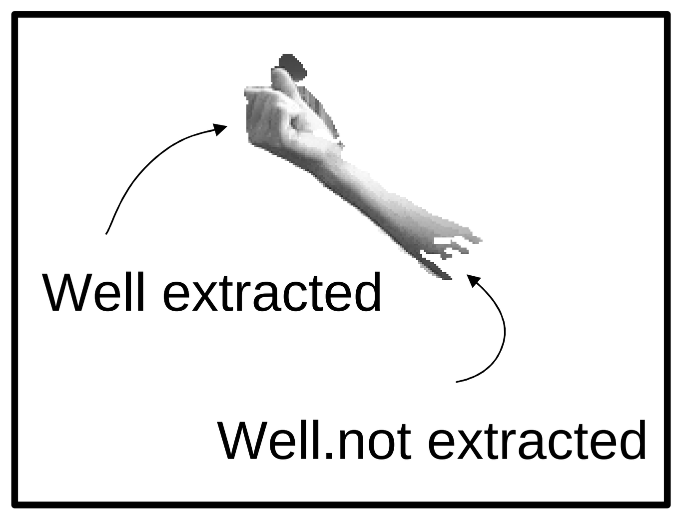
EXPERIMENTAL RESULTS

Original image

Indirect method



Direct method

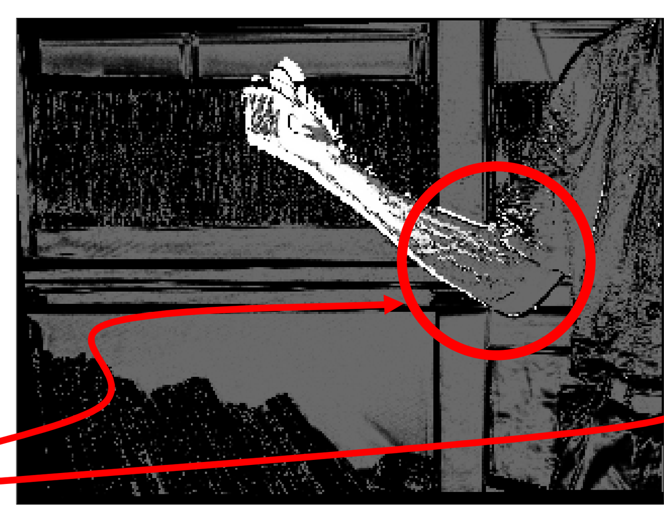


Posterior Probability

Black : low prob.
White : high prob.

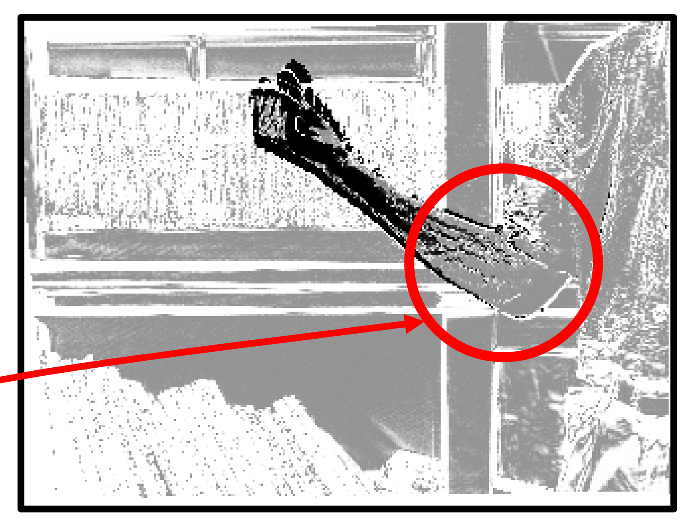
gray : ambiguous

Arm region



$P(\theta_1)P(r_j\,|\,\theta_1)$

Background



$P(\theta_0)P(r_j\,|\,\theta_0)$

Original image | **Extraction of Indirect method** | **Extraction of combined method**