Intelligent Control Systems

Visual Tracking (2)
— Histogram-based Tracking with Mean Shift Method —

Shingo Kagami
Graduate School of Information Sciences,
Tohoku University
swk(at)ic.is.tohoku.ac.jp

http://www.ic.is.tohoku.ac.jp/ja/swk/

Histogram-based Tracking

- Lucas-Kanade method, which we introduced last week, is suited for tracking of rigid objects
  - It is possible to adapt to object deformation by updating the template (as we demonstrated), then the risk of drifting away from the original target becomes greater

- Are there any features that are invariant with shape deformation?
  - Color histogram
  - Mean Shift method [Fukunaga 1975] applied to tracking of a region with color histogram similar to a given model histogram [Comaniciu 2003]
    - Note: this is different from CAMSHIFT algorithm available as `cv::meanShift()`

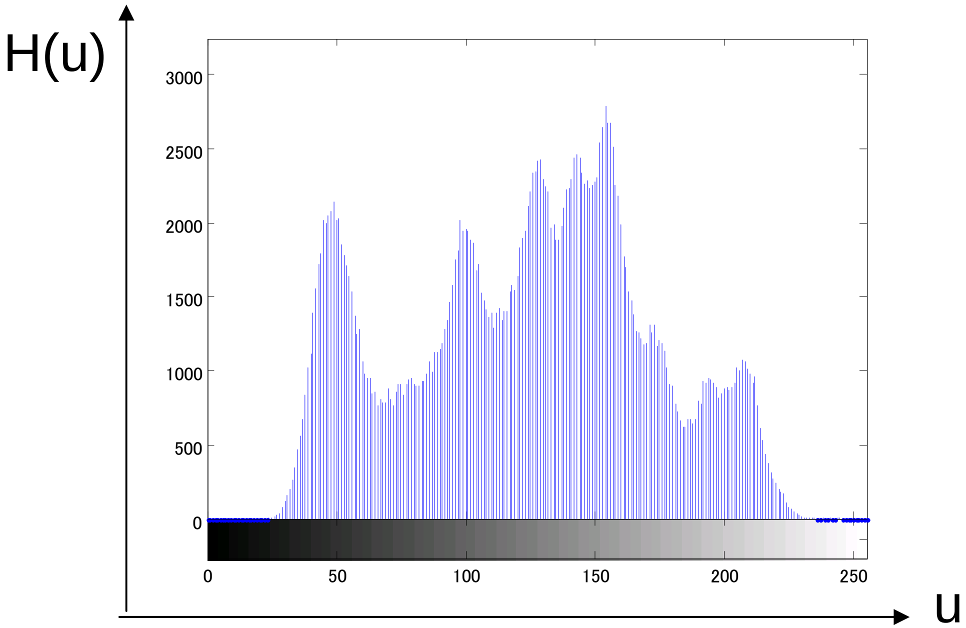
Outline

- Color Histograms
  - weighted histograms
  - similarity measure of histograms
  - approximation of the similarity measure
  (to make the problem dealt with by mean shift)

- Mean Shift Method

- Mean Shift Tracking based on Color Histogram

(Grayscale) Histogram



$$\boldsymbol{H} = \{H_u\}_{u=1,2,\cdots,m}, \; H_u = \sum_{\boldsymbol{x} \in S(u)} 1$$

where S(u) is a set of pixels having values belonging to the bin u

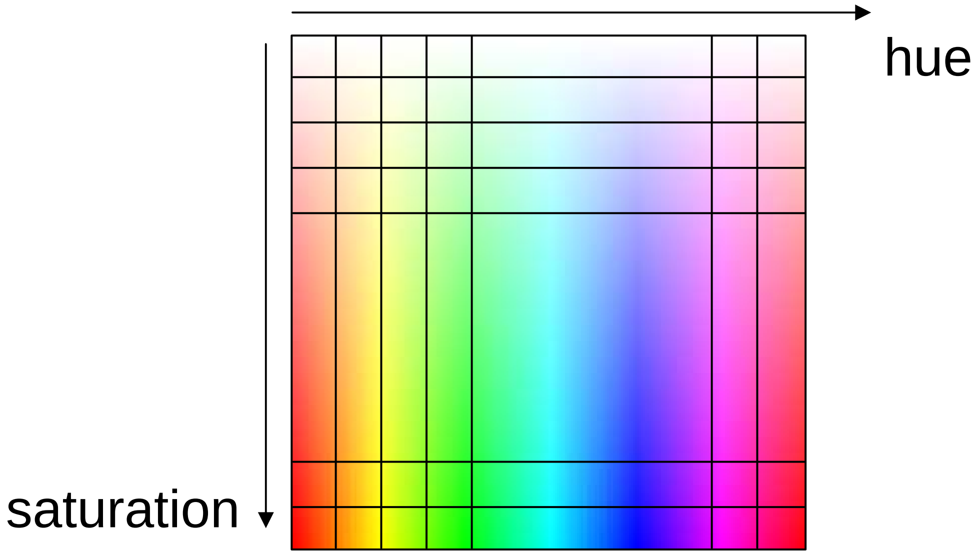
$$\boldsymbol{p} = \{p_u\}, \; p_u \propto H_u, \; \sum_{u=1}^{m} p_u = 1$$ (normalized histogram)

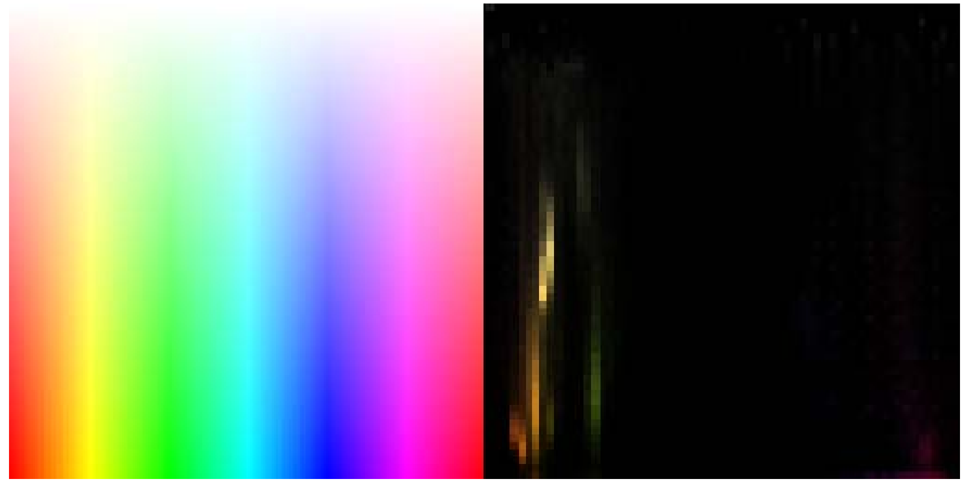
Color Histograms

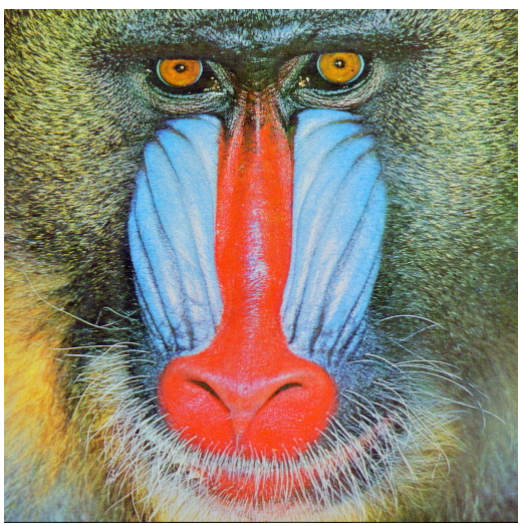
ex1) By splitting each of RGB components into 16 bins, we have histogram over 16 × 16 × 16 bins

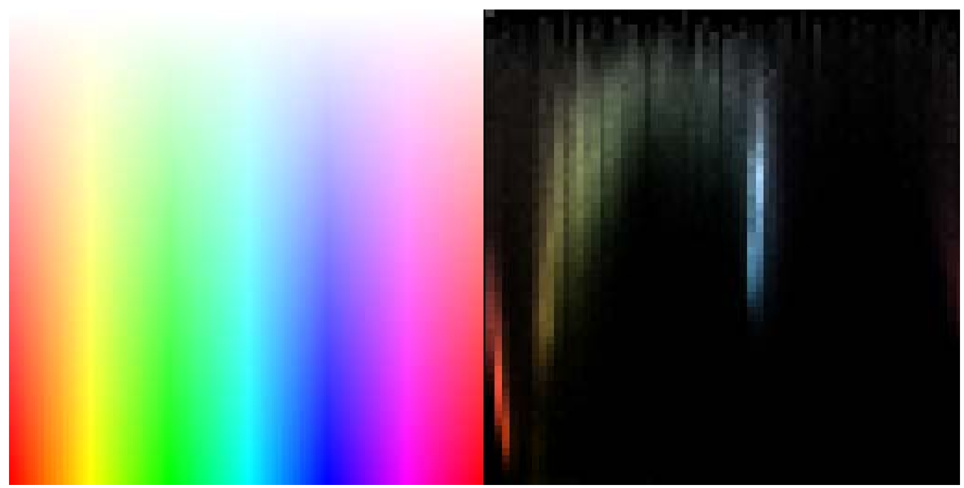
ex2) By splitting each of Hue and Saturation components into 64 bins (and ignoring Value component), we have histogram over 64 × 64 bins

- More unaffected by illumination change









Similarity of Histograms

- Our objective is to find a region with histogram similar to that of a given model
- How do we measure the similarity?

Bhattacharyya Coefficient

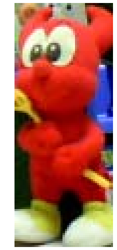
- is a metric for similarity of two probabilistic distributions (and thus, of two normalized histograms) **p** and **q**

$$\rho(\boldsymbol{p}, \boldsymbol{q}) = \sum_{u=1}^{m} \sqrt{p_u q_u}$$

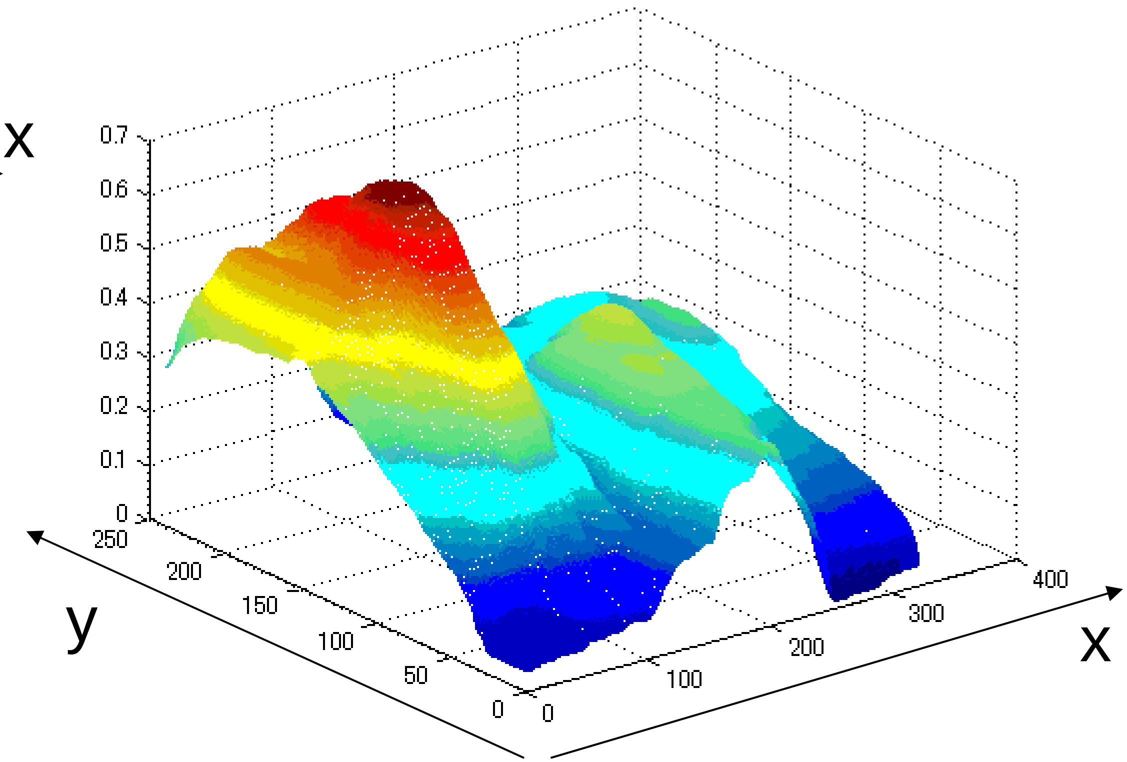
- Geometric interpretation: inner product of $(\sqrt{q_1}, \sqrt{q_2}, \cdots, \sqrt{q_m})^{\mathsf{T}}$ and $(\sqrt{p_1}, \sqrt{p_2}, \cdots, \sqrt{p_m})^{\mathsf{T}}$, which lie on the unit sphere surface

Similarity Map

current image



object model



Weighted Histogram

- The pixels near boundaries should have small influence
- Discontinuity in the similarity map is not favored

$\rightarrow$ **weight the voting depending of pixel locations**

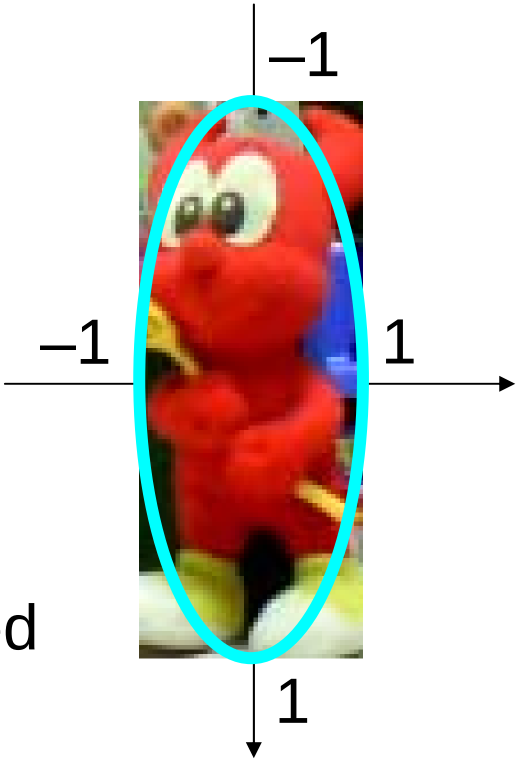
Object Model:

$$q_u \propto \sum_{x \in S_0(u)} k(||x||^2)$$

$S_0(u)$: Set of pixels whose pixel values belong to bin u in the model image

k(): weight function or kernel function

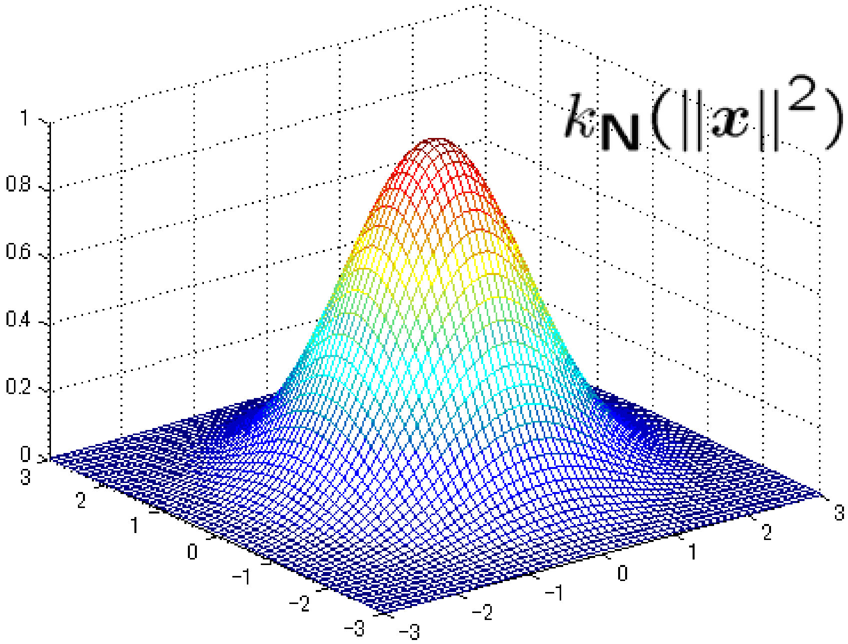
- centered at origin
- image coordinates **x** = (x, y) are normalized so that it fits the unit circle



Kernel Function Examples

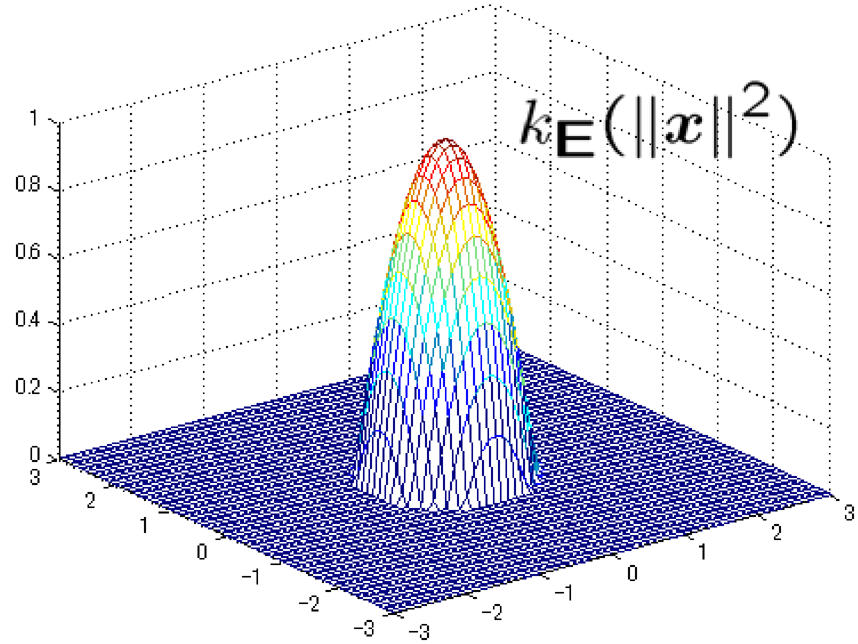
ex1) Gauss kernel

$$k_{\mathbf{N}}(x) \propto \exp(-\frac{x}{2})$$



ex2) Epanechnikov kernel
(kernel with Epanechnikov profile)

$$k_{\mathbf{E}}(x) \propto \begin{cases} 1 - x, & x \leq 1 \\ 0, & \textbf{otherwise} \end{cases}$$



Weighted Histogram of Candidate Region

Histogram of candidate region (centered at **y**):

$$p_u(\boldsymbol{y}) \propto \sum_{\boldsymbol{x} \in S(u)} k(\|\boldsymbol{y} - \boldsymbol{x}\|^2)$$

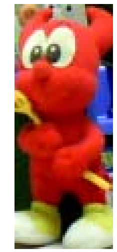
unit circle centered at **y**

S(u): Set of pixels whose pixel values belong to bin u in the current image

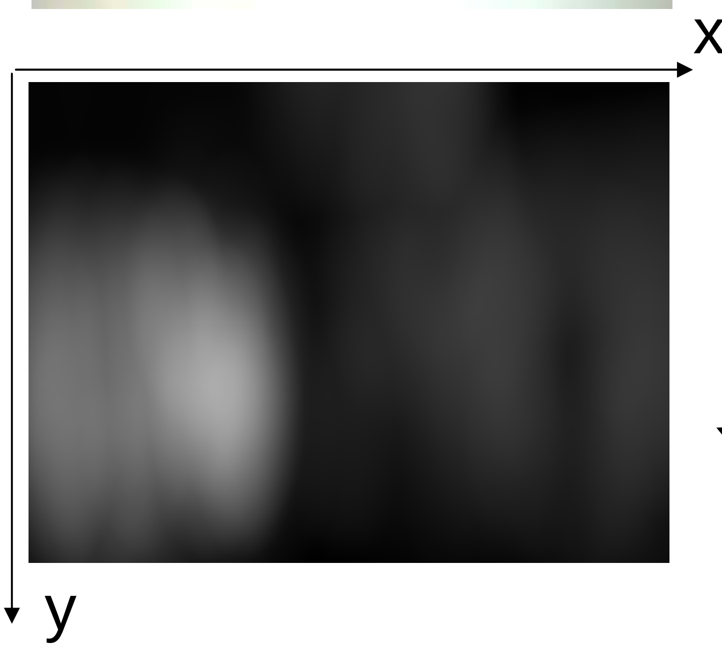
- Note again: image coordinates are normalized so that it fits the unit circle

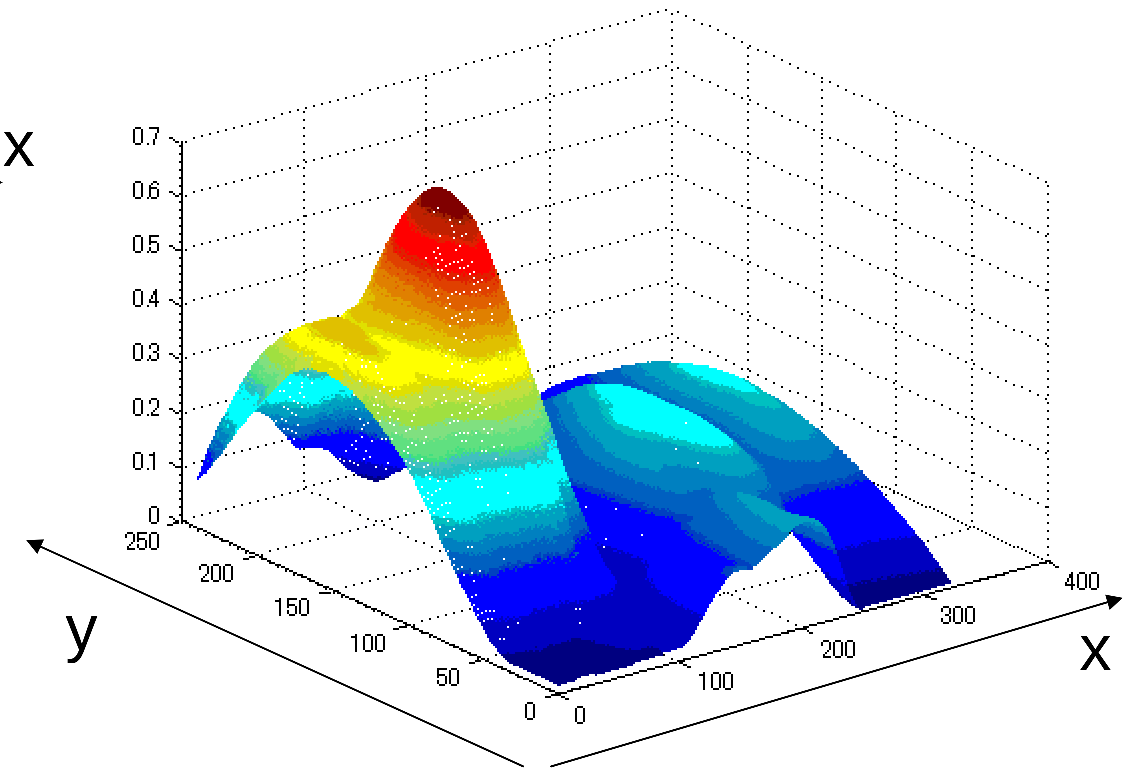
Similarity Map with Weighted Histogram

current image



object model





Outline

- Color Histograms
  - weighted histograms
  - similarity measure of histograms
  - approximation of the similarity measure
  (to make the problem dealt with by mean shift)

- Mean Shift Method

- Mean Shift Tracking based on Color Histogram

Approximating the Similarity

Since exhaustive search for maximum similarity is too time consuming, let's think of using gradient methods

- Let the initial candidate position be $\mathbf{y_0}$
- Consider 1st order Taylor expansion to $\rho(\mathbf{p}(\mathbf{y}), \mathbf{q})$ with respect to $\mathbf{p}(\mathbf{y})$ around $\mathbf{p}(\mathbf{y}_0)$

$$\begin{aligned}
\rho(\boldsymbol{p}(\boldsymbol{y}), \boldsymbol{q}) &= \sum_u \sqrt{p_u(\boldsymbol{y})}\sqrt{q_u} \\
&\approx \sum_u \sqrt{q_u}\left\{\sqrt{p_u(\boldsymbol{y}_0)} + \frac{1}{2}p_u(\boldsymbol{y}_0)^{-1/2}\left(p_u(\boldsymbol{y}) - p_u(\boldsymbol{y}_0)\right)\right\} \\
&= \sum_u \sqrt{q_u}\left(\sqrt{p_u(\boldsymbol{y}_0)} + \frac{1}{2}p_u(\boldsymbol{y})\frac{1}{\sqrt{p_u(\boldsymbol{y}_0)}} - \frac{1}{2}\sqrt{p_u(\boldsymbol{y}_0)}\right) \\
&= \frac{1}{2}\sum_u \sqrt{q_u}\sqrt{p_u(\boldsymbol{y}_0)} + \frac{1}{2}\sum_u p_u(\boldsymbol{y})\frac{\sqrt{q_u}}{\sqrt{p_u(\boldsymbol{y}_0)}}
\end{aligned}$$

Since the 1st term does not depend on y, what we should maximize is the 2nd term:

$$\sum_u p_u(\boldsymbol{y}) \frac{\sqrt{q_u}}{\sqrt{p_u(\boldsymbol{y}_0)}}$$

Recalling that $p_u(\boldsymbol{y}) \propto \sum_{\boldsymbol{x} \in S(u)} k(\|\boldsymbol{y} - \boldsymbol{x}\|^2)$,

this comes down to maximization of

$$\sum_{u \in \textbf{all bins}} \; \sum_{x \in \textbf{all pixels belonging to } u} k(\|\boldsymbol{y} - \boldsymbol{x}\|^2) \frac{\sqrt{q_u}}{\sqrt{p_u(\boldsymbol{y}_0)}}$$

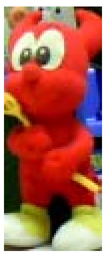
$$= \sum_{x \in \textbf{all pixels}} \sqrt{\frac{q_{b(x)}}{p_{b(x)}(\boldsymbol{y}_0)}} k(\|\boldsymbol{y} - \boldsymbol{x}\|^2)$$

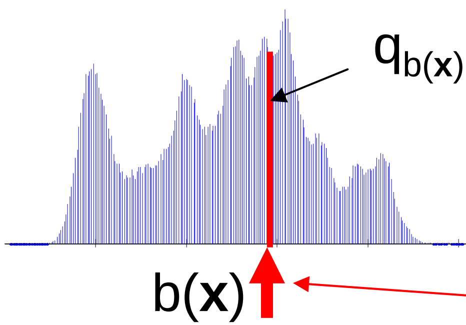
where b(**x**) is the bin to which x belongs

So, what we should maximize is:

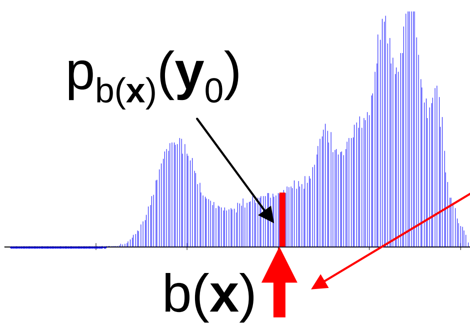
$$\sum_{x \in \text{all pixels}} \sqrt{\frac{q_{b(x)}}{p_{b(x)}(y_0)}}\, k(\|y - x\|^2) = \sum_{x \in \text{all pixels}} w(x) k(\|y - x\|^2)$$

model histogram





histogram of region around $\mathbf{y}_0$



For each pixel **x** (in the kernel range), find b(**x**) to look up q and p, and compute w(x).

Outline

- Color Histograms
  - weighted histograms
  - similarity measure of histograms
  - approximation of the similarity measure
  (to make the problem dealt with by mean shift)

- Mean Shift Method

- Mean Shift Tracking based on Color Histogram

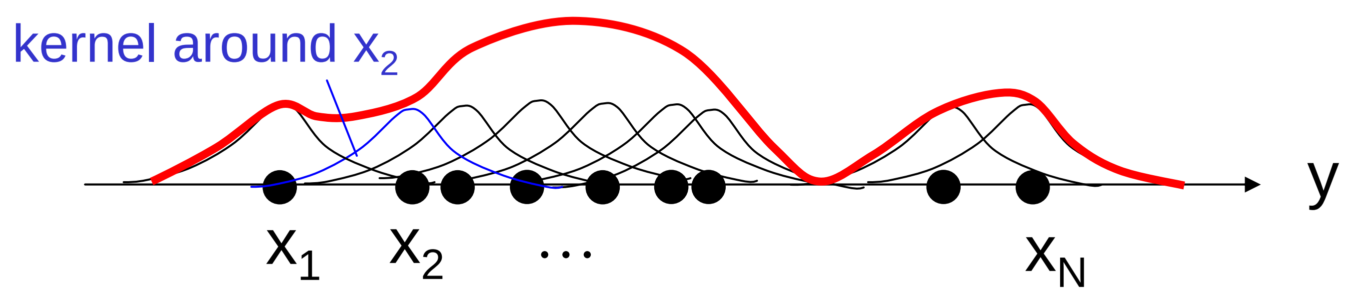
Kernel Density Estimation

$$\sum_{x \in \text{all pixels}} w(x) k(\|y - x\|^2)$$

- Our view until now: k() lies around **y** (and evaluated at each **x**, and summed over **x**)
- New view: k() is around each **x** (and evaluated at a single point **y**, and summed over **x**)

This new view is often used when one wants to estimate a probability distribution from finite samples drawn from the distribution

- called Kernel Density Estimation (KDE) or Parzen Estimation



Mean Shift Method

- An efficient method to find a local maximum of a probability distribution estimated by KDE

$$f_k(\boldsymbol{y}) = \sum_{\boldsymbol{x}} w(\boldsymbol{x}) k(\|\boldsymbol{y} - \boldsymbol{x}\|^2)$$

Gradient: $\nabla f_k(\boldsymbol{y}) = \frac{\partial}{\partial \boldsymbol{y}} f_k(\boldsymbol{y}) = \sum_{\boldsymbol{x}} k'(\|\boldsymbol{y} - \boldsymbol{x}\|^2) \cdot 2(\boldsymbol{y} - \boldsymbol{x}) w(\boldsymbol{x})$

Writing $g(x) = -k'(x)$, we have

$$\nabla f_k(\boldsymbol{y}) = 2 \sum_{\boldsymbol{x}} g(\|\boldsymbol{y} - \boldsymbol{x}\|^2)(\boldsymbol{x} - \boldsymbol{y}) w(\boldsymbol{x})$$

KDE with kernel g

$$= 2 \left[\sum_{\boldsymbol{x}} \left\{ \boldsymbol{x} w(\boldsymbol{x}) g(\|\boldsymbol{y} - \boldsymbol{x}\|^2) \right\} - \boldsymbol{y} \sum_{\boldsymbol{x}} \left\{ w(\boldsymbol{x}) g(\|\boldsymbol{y} - \boldsymbol{x}\|^2) \right\} \right]$$

$$= 2 f_g(\boldsymbol{y}) \left[\frac{\sum_{\boldsymbol{x}} \left\{ \boldsymbol{x} w(\boldsymbol{x}) g(\|\boldsymbol{y} - \boldsymbol{x}\|^2) \right\}}{f_g(\boldsymbol{y})} - \boldsymbol{y} \right]$$

Mean Shift vector:

$$m_g(y) = \frac{\nabla f_k(y)}{2 f_g(y)} = \left[\frac{\sum_x \left\{ x w(x) g(\|y - x\|^2) \right\}}{\sum_x \left\{ w(x) g(\|y - x\|^2) \right\}} - y \right]$$

- toward the direction $f_k(\mathbf{y})$ becomes larger
- large when $f_g(\mathbf{y})$ is small, small when $f_g(\mathbf{y})$ is large

As a special case, if you use Epanechnikov kernel as k(), g() becomes 1 within the unit circle, and 0 otherwise

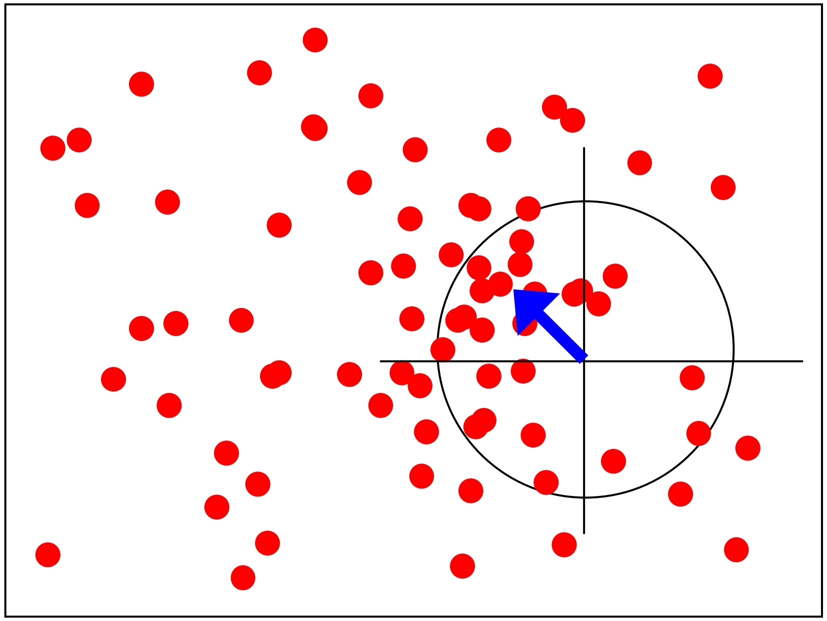
$$f_g(y) = \sum_{x \in \textbf{unit circle}} w(x)$$

center of gravity within unit circle

$$m_g(y) = \left[\frac{\sum_{x \in \textbf{unit circle}} x w(x)}{\sum_{x \in \textbf{unit circle}} w(x)} - y \right]$$

Mean Shift Method (general KDE problem)

1. Compute center of gravity of samples around current position
2. Move to the center of gravity (Mean Shift)
3. Return to 1. unless the mean shift vector is too small



- When the maximum is far, $f_g(\mathbf{y})$ will be small and the shift will be large
- When the maximum is near, the shift will become small
- It has been proved that it converges to the local maximum with mild conditions for the shape of k()
  - Epanechnikov kernel satisfies it

Outline

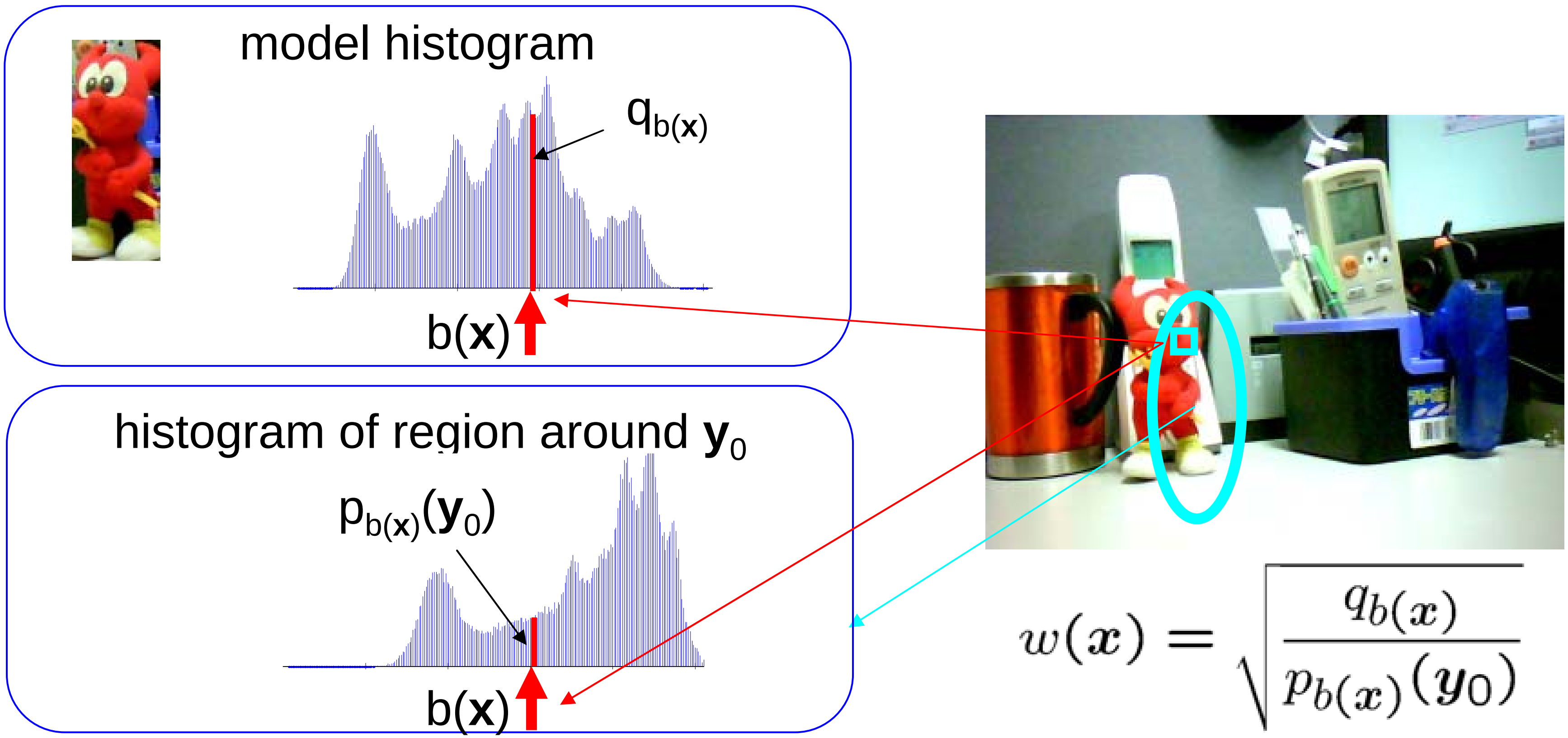
- Color Histograms
  - weighted histograms
  - similarity measure of histograms
  - approximation of the similarity measure
    (to make the problem dealt with by mean shift)

- Mean Shift Method

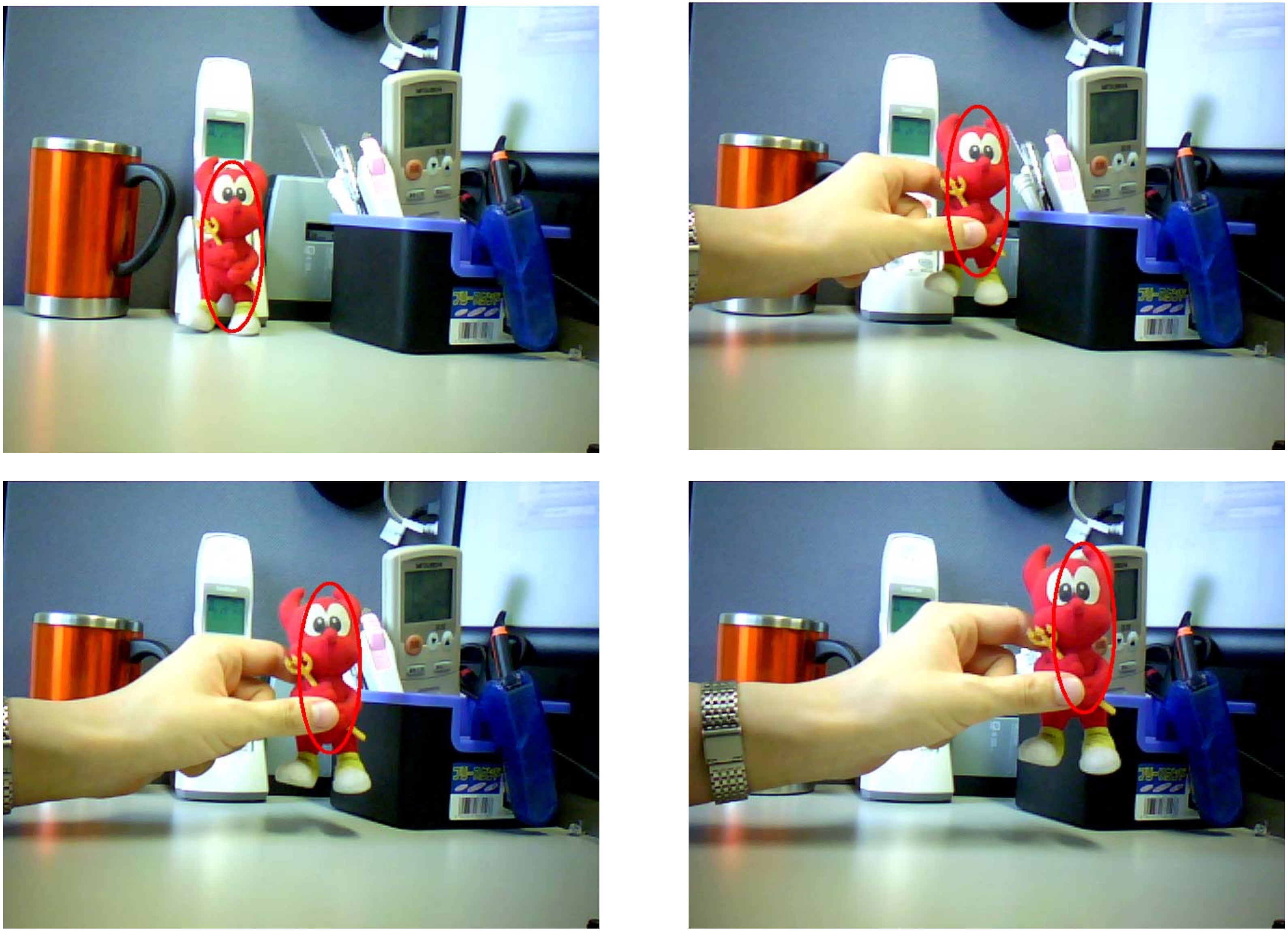
- Mean Shift Tracking based on Color Histogram

Mean Shift Method (for histogram tracking)

1. Compute the weighted histogram p($\mathbf{y}_0$) around $\mathbf{y}_0$
2. Move $\mathbf{y}_0$ to the center of gravity of w($\mathbf{x}$), by finding b($\mathbf{x}$) and looking up **q** an **p** for each pixel **x** around $\mathbf{y}_0$
3. Return to 1. unless the move is too small



$$w(x) = \sqrt{\frac{q_{b(x)}}{p_{b(x)}(y_0)}}$$



Summary

- Color Histograms
  - can be defined over any color spaces
  - Hue-Saturation color space is often preferred due to its invariance with respect to illumination
- Weighted Histograms
  - avoid boundary effects and discontinuity in similarity map
  - make mean shift applicable
    - (Epanechnikov kernel)’ = unit circle
- Mean Shift
  - finds local maximum of a probability distribution estimated by kernel density estimation (KDE)
  - Similarity map of histograms (by Bhattacharyya coefficient) can be viewed as KDE

References

- D. Comaniciu, V. Ramesh and P. Meer: Kernel-Based Object Tracking, IEEE Trans. of Pattern Analysis and Machine Intelligence, vol.25, no.5, 2003.
- D. Comaniciu and P. Meer: Mean Shift: A Robust Approach Toward Feature Space Analysis, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.25, no.5, 2003.
- K. Fukunaga and L. D. Hostetler: The Estimation of the Gradient of a Density Function, with Applications in Pattern Recognition, IEEE Trans. on Information Theory, vol.IT-21, no.1, 1975.

Sample codes are in sample20140708.zip available at
http://www.ic.is.tohoku.ac.jp/~swk/lecture/