Saliency-Based Image Editing for Guiding Visual Attention

Aiko Hagiwara (Chiba University, JAPAN)

Sugimoto Akihiro (NII, JAPAN)

Kazuhiko Kawamoto (Chiba University, JAPAN)

大学共同利用機関法人 情報・システム研究機構
NII 国立情報学研究所
National Institute of Informatics

Guiding visual attention
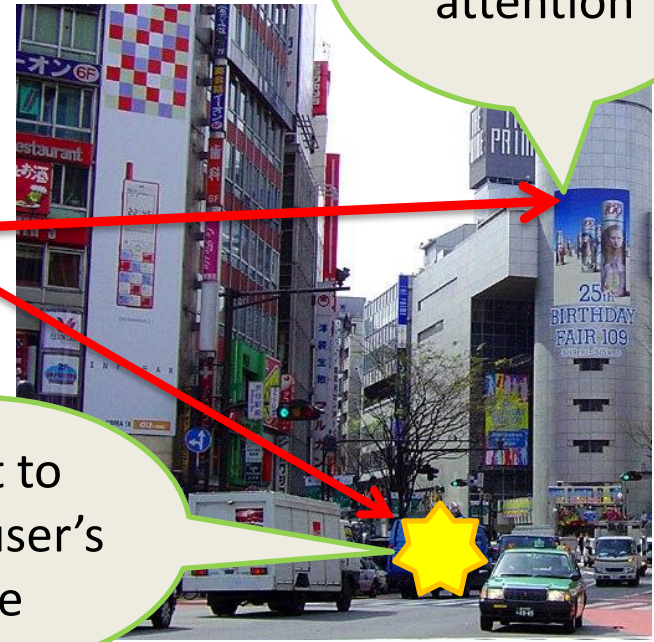
Presenting external visual stimulus

↓

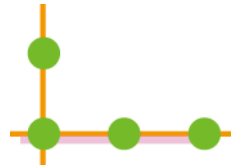Naturally and smoothly, without interrupting current gaze.
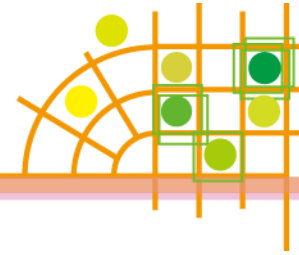
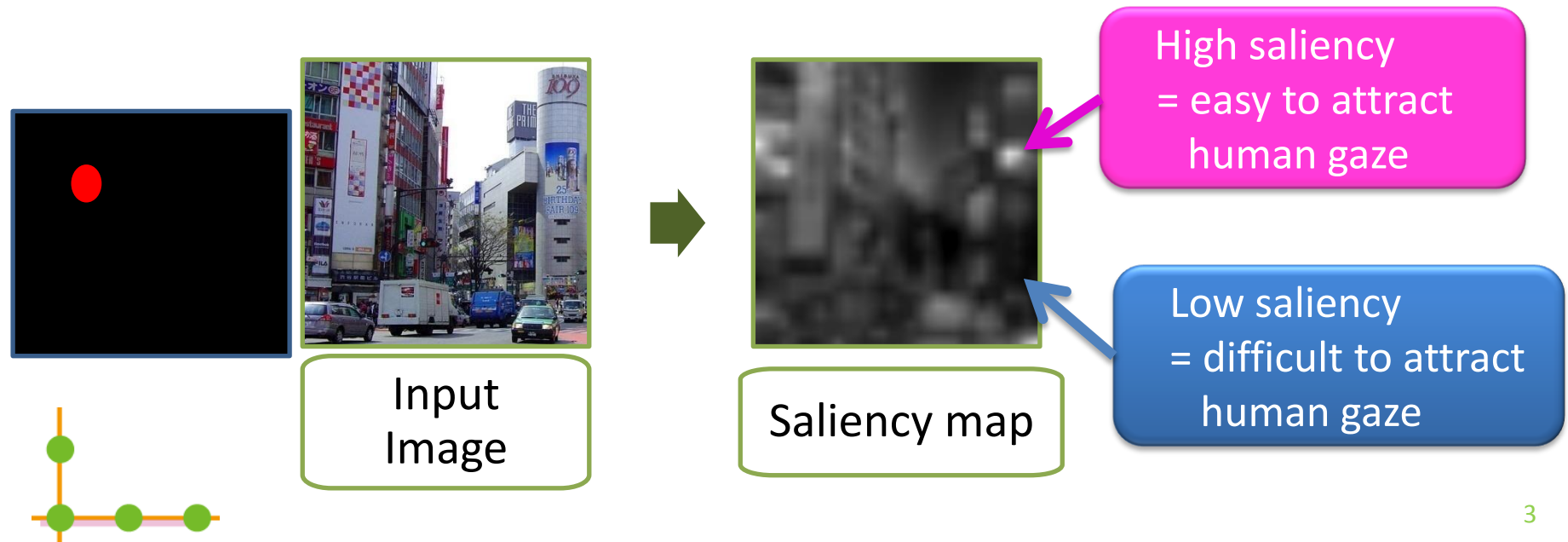Attract user's visual attention

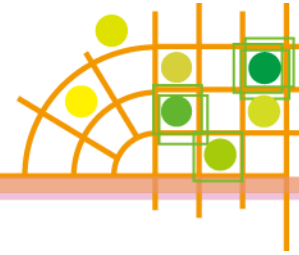Want to guide user's gaze

© 2011 BIZUTART PHOTORAPH

Visual attention

- Gaze point selection is controlled by…

  Human's interest   & Image itself

- Computational model of saliency using bottom-up
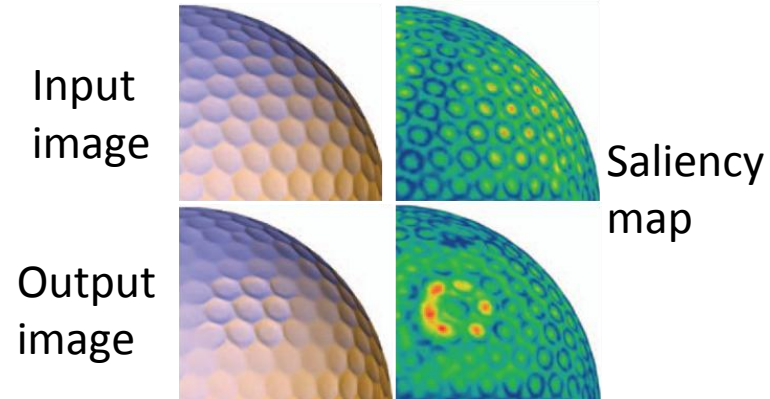
  information obtained from image. [Itti,et al. 1998]

Input Image

Saliency map

High saliency = easy to attract human gaze

Low saliency = difficult to attract human gaze

3

Related work

- ## Subtle Gaze Direction[Baiely,et al. 2009]

  Presenting visual stimulus in the user's peripheral vision area.

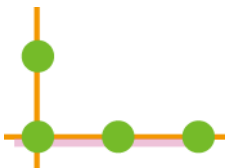Input image   Stimulus

- ## Persuading visual attention through geometry[Kim,et al. 2008]

  Introducing geometry modification.

Input image

Saliency map

Output image

Directly modulating saliency

Computational model of saliency map

Input image

extract

Intensity i Color r g b y

make Gaussian pyramids

make feature maps using across-scale difference

rg by

$\overline{i}$ fuse $\overline{c}$
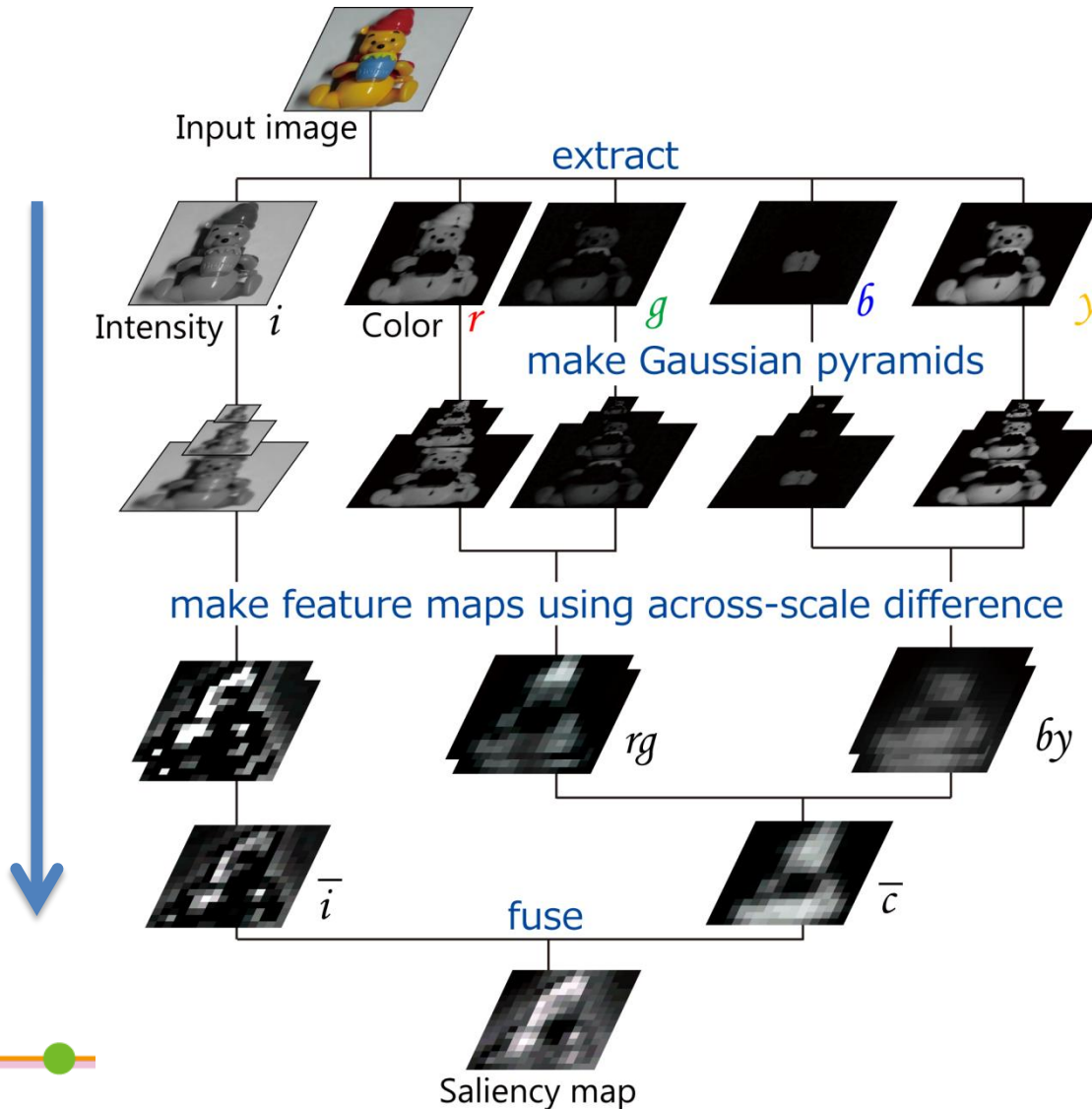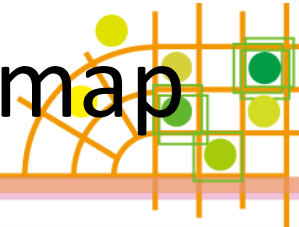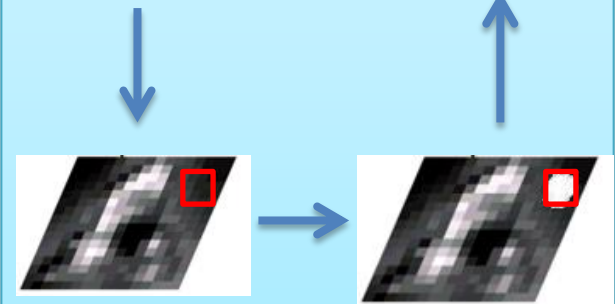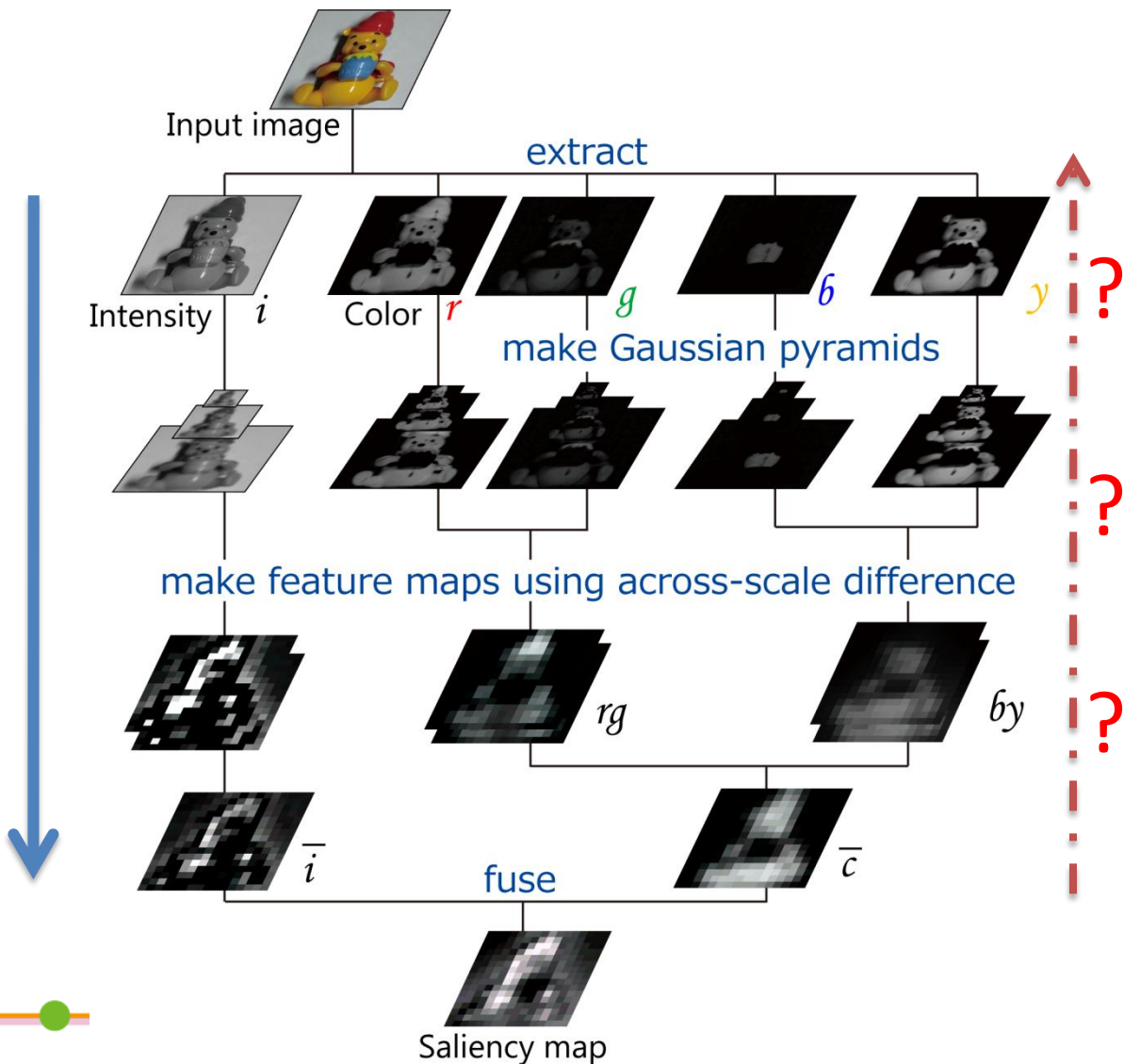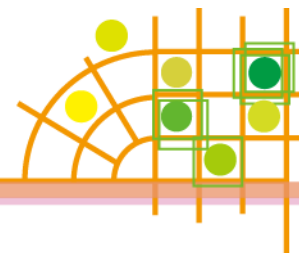
Saliency map

Our problem

Given region = where you want to guide user's gaze

raise given region's saliency

?

Our approach

Input image
extract
Intensity i Color r g b y
make Gaussian pyramids
make feature maps using across-scale difference
rg by
$\overline{i}$ fuse $\overline{c}$
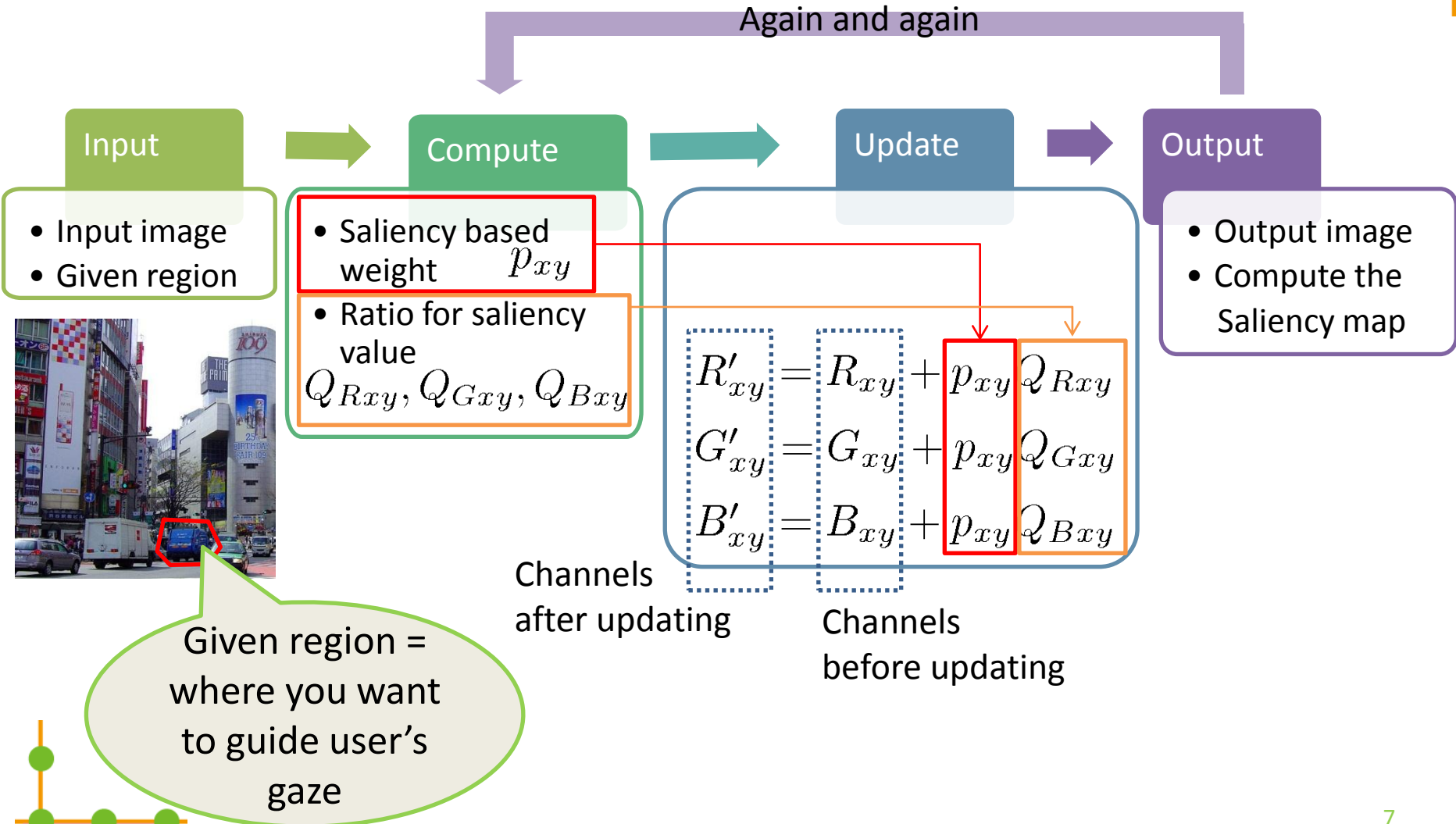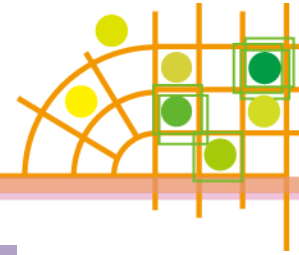Saliency map

Inside the region = Raise saliency = strengthen the feature

**Outside the region =Reduce saliency =weaken the feature**

6

Proposed method

Again and again

Input	Compute	Update	Output

Input
- Input image
- Given region

Compute
- Saliency based weight p_{xy}
- Ratio for saliency value
  $Q_{Rxy}, Q_{Gxy}, Q_{Bxy}$

Update

$$R'_{xy} = R_{xy} + p_{xy} Q_{Rxy}$$
$$G'_{xy} = G_{xy} + p_{xy} Q_{Gxy}$$
$$B'_{xy} = B_{xy} + p_{xy} Q_{Bxy}$$

Channels after updating

Channels before updating

Output
- Output image
- Compute the Saliency map

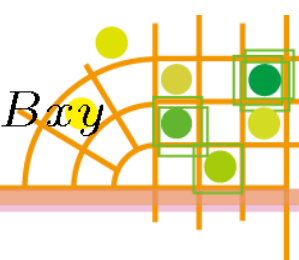Given region = where you want to guide user's gaze
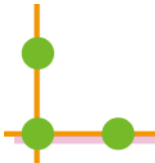
Saliency based weight p_{xy}
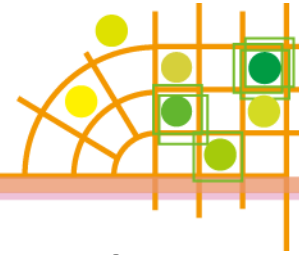
- Magnitude: proportional to saliency value at that point

- Sign: positive inside, negative outside

Strengthen

Given region

p_{xy}

weaken
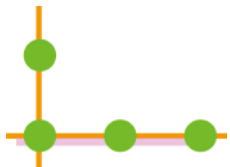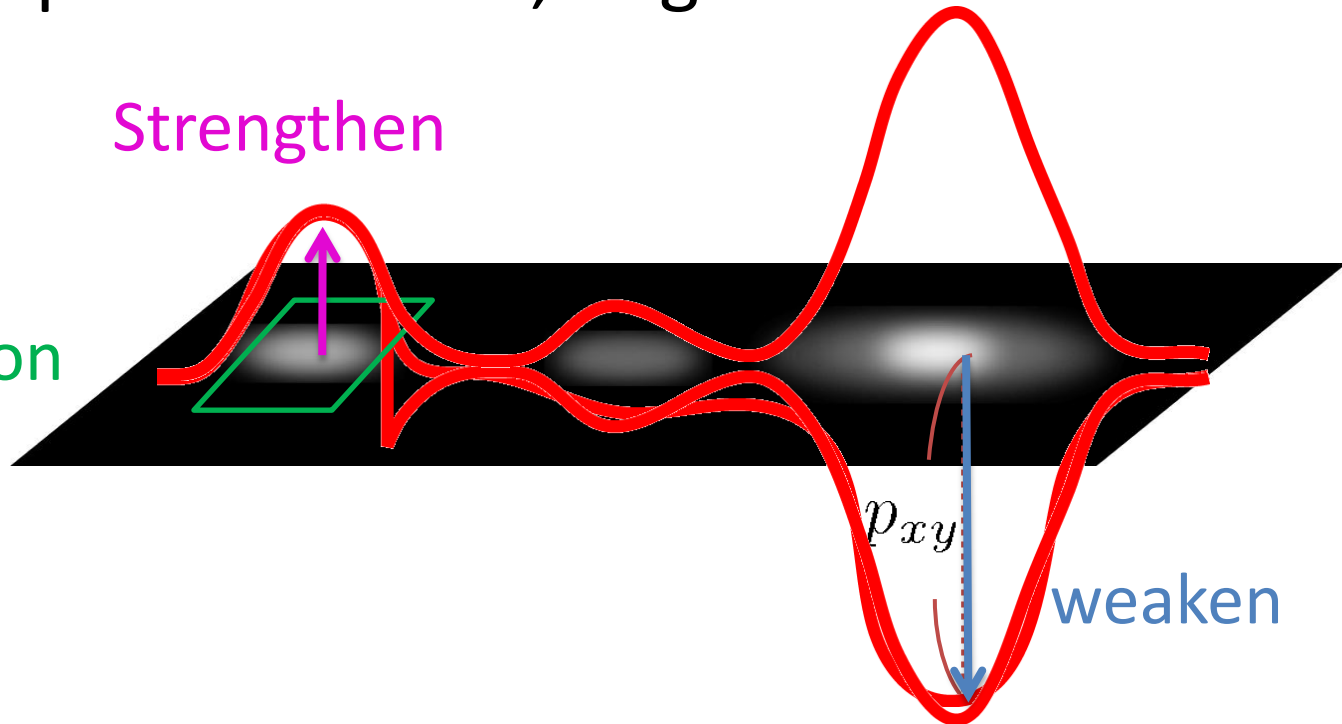
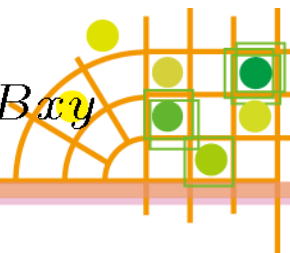# Ratio for saliency value $Q_{Rxy}, Q_{Gxy}, Q_{Bxy}$

- Computing influence ratio for each component of color space(i,r,g,b,y)

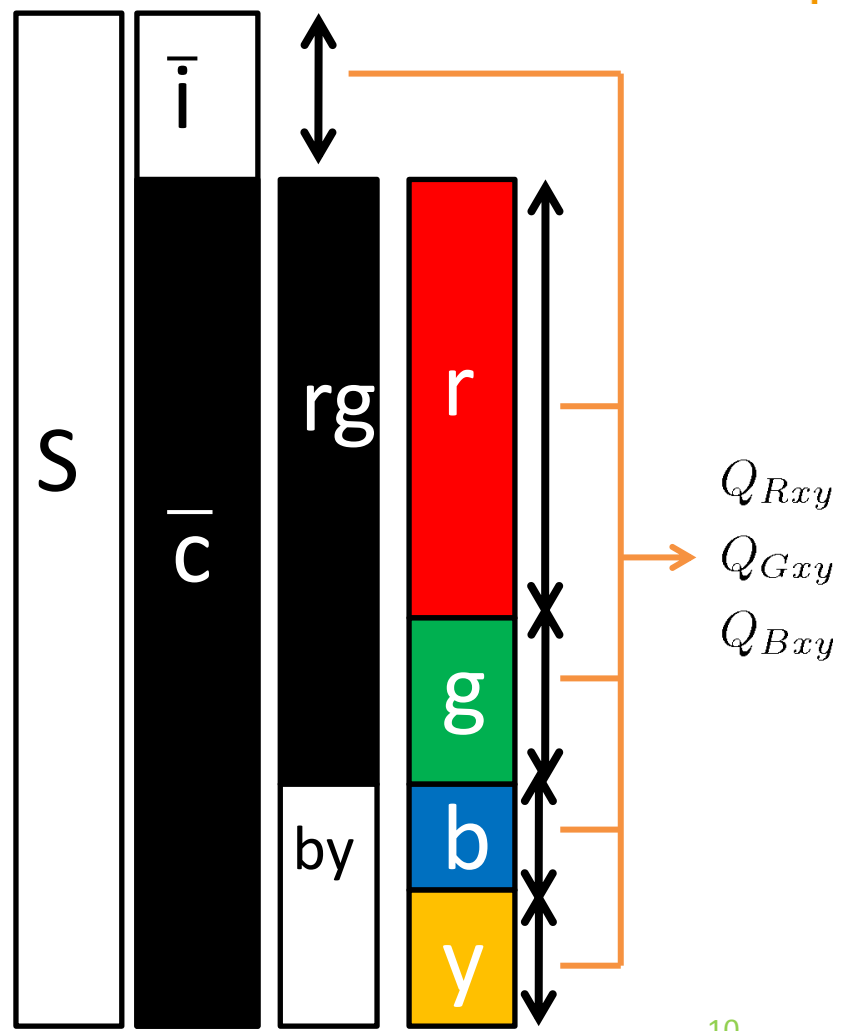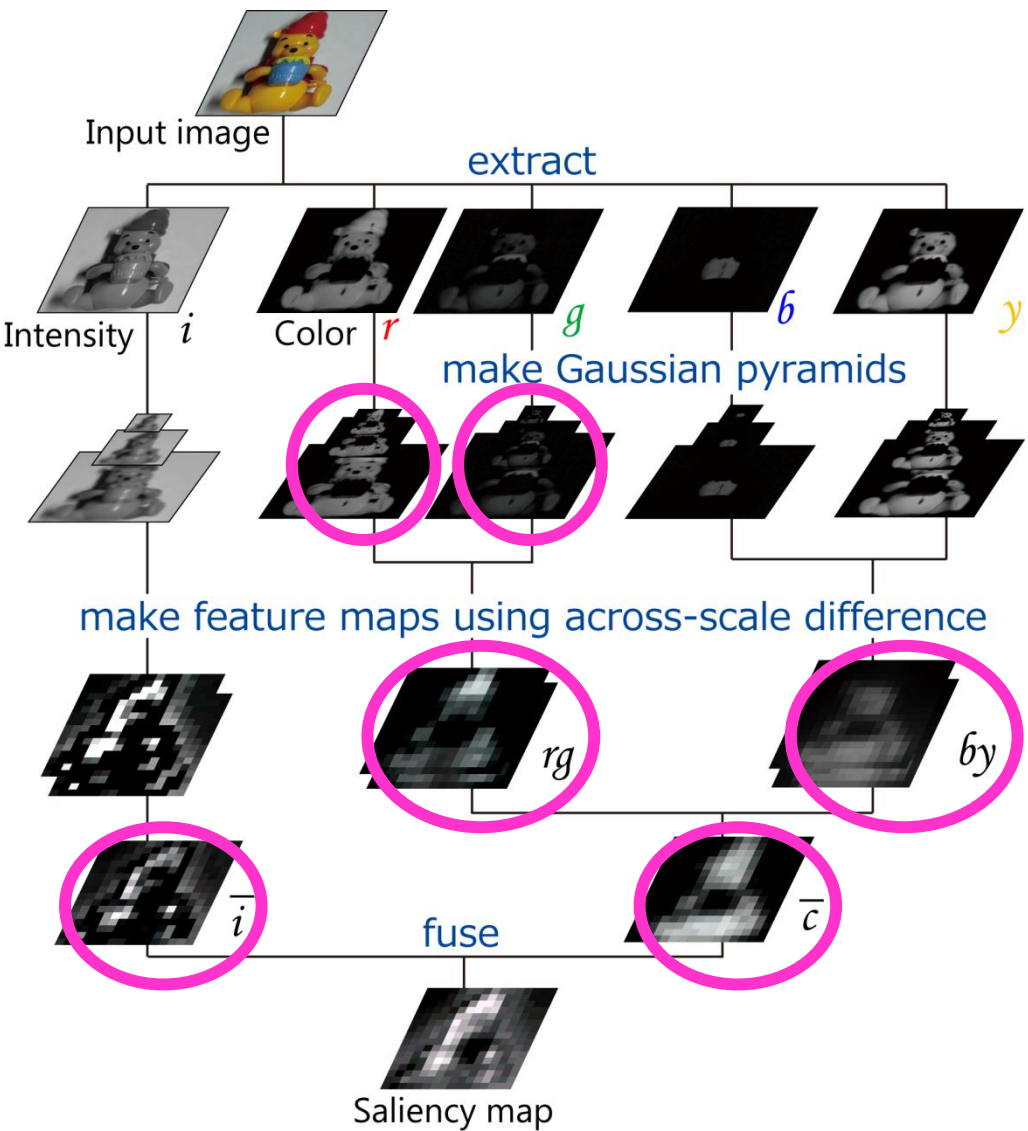 →adjusting R,G,B using these ratios.

It is salient because of "Red"
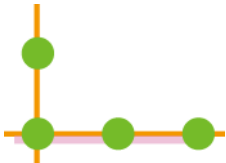
Q_{Rxy} large

Q_{Gxy} small

Q_{Bxy} small

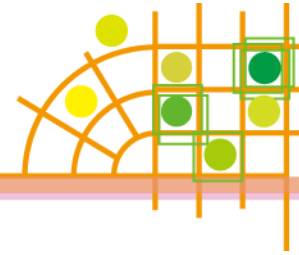# Ratio for saliency value $Q_{Rxy}, Q_{Gxy}, Q_{Bxy}$

Input image

extract

Intensity i Color r g b y

make Gaussian pyramids

make feature maps using across-scale difference

rg by

$\bar{i}$ fuse $\bar{c}$

Saliency map

S $\bar{i}$ $\bar{c}$ rg   by   r   g   b   y

Q_{Rxy}
Q_{Gxy}
Q_{Bxy}

10

Proposed method

Again and again

Input
- Input image
- Given region

Compute
- Saliency based weight p_{xy}
- Ratio for saliency value $Q_{Rxy}, Q_{Gxy}, Q_{Bxy}$

Update

$$R'_{xy} = R_{xy} + p_{xy} Q_{Rxy}$$
$$G'_{xy} = G_{xy} + p_{xy} Q_{Gxy}$$
$$B'_{xy} = B_{xy} + p_{xy} Q_{Bxy}$$

Channels after updating

Channels before updating

Output
- Output image
- Compute the Saliency map

Experiment

Image editing result

Image editing result

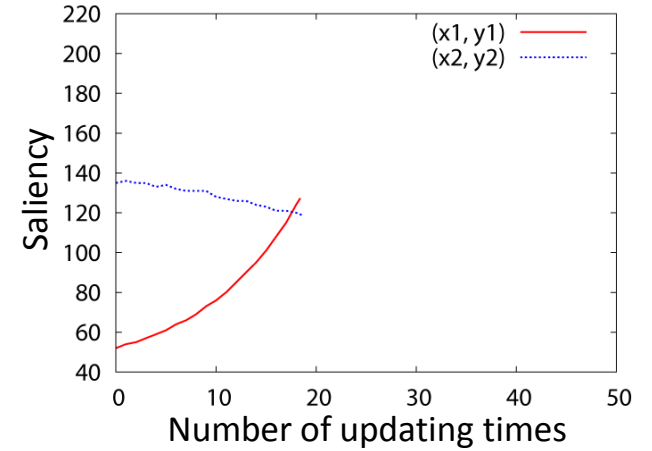

↓Saliency map
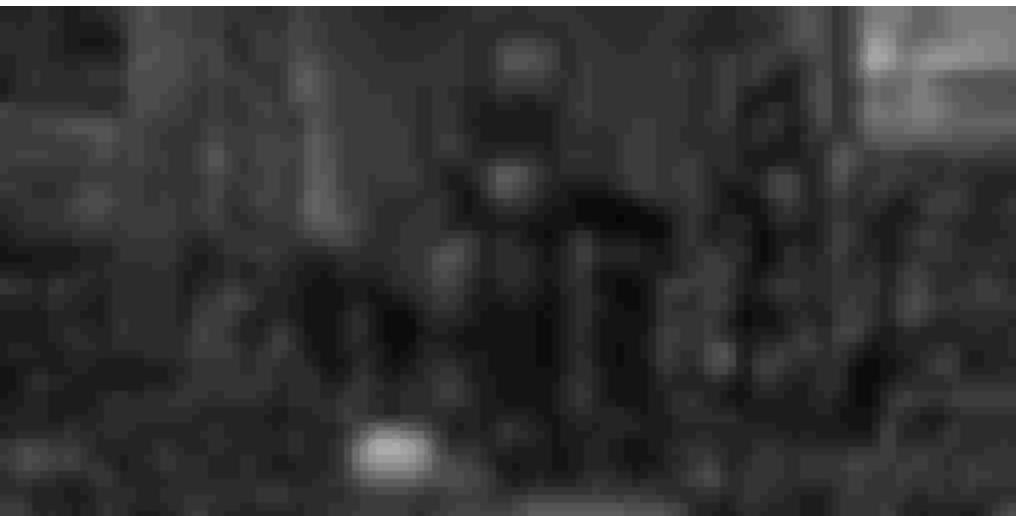
Saliency value
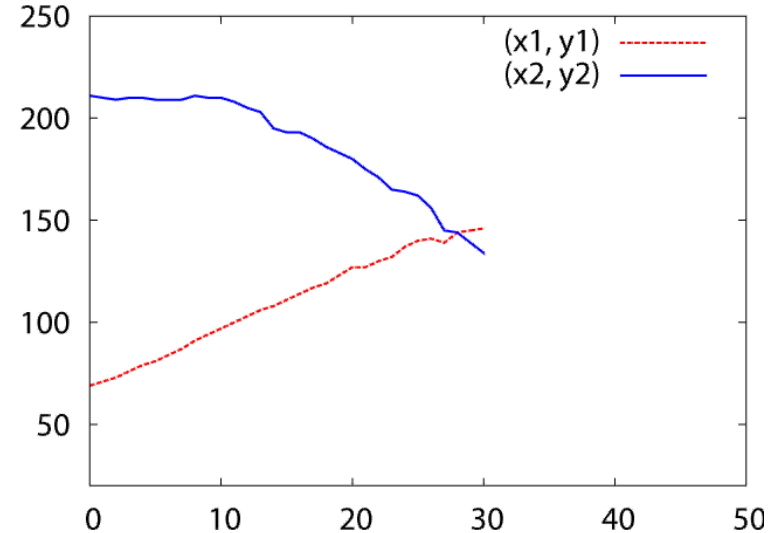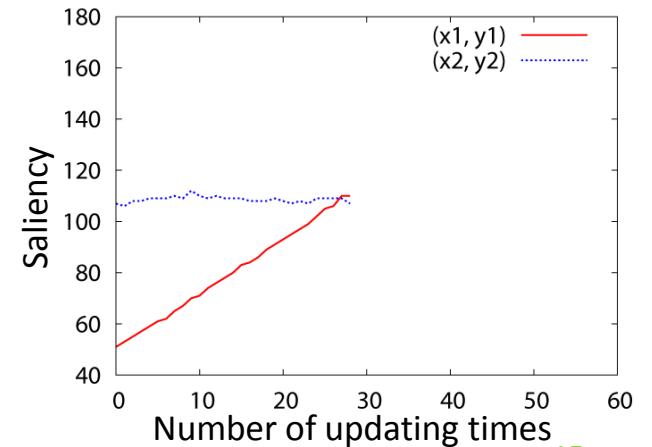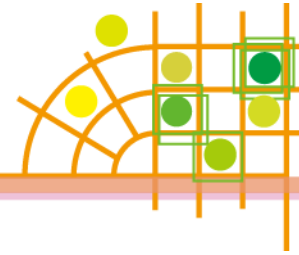

Number of updating times

14

Image editing results
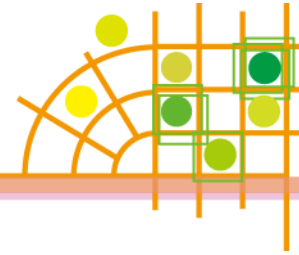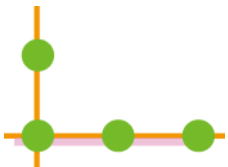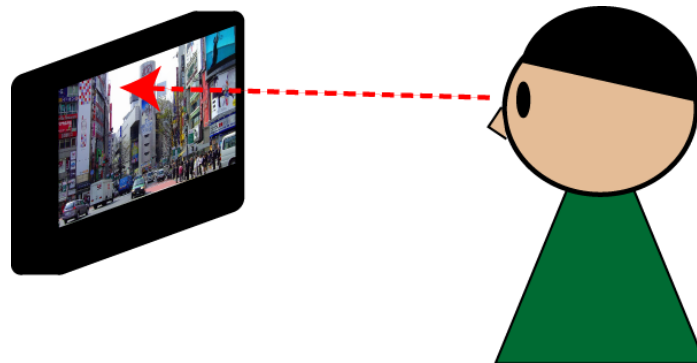

Original image | Edited image
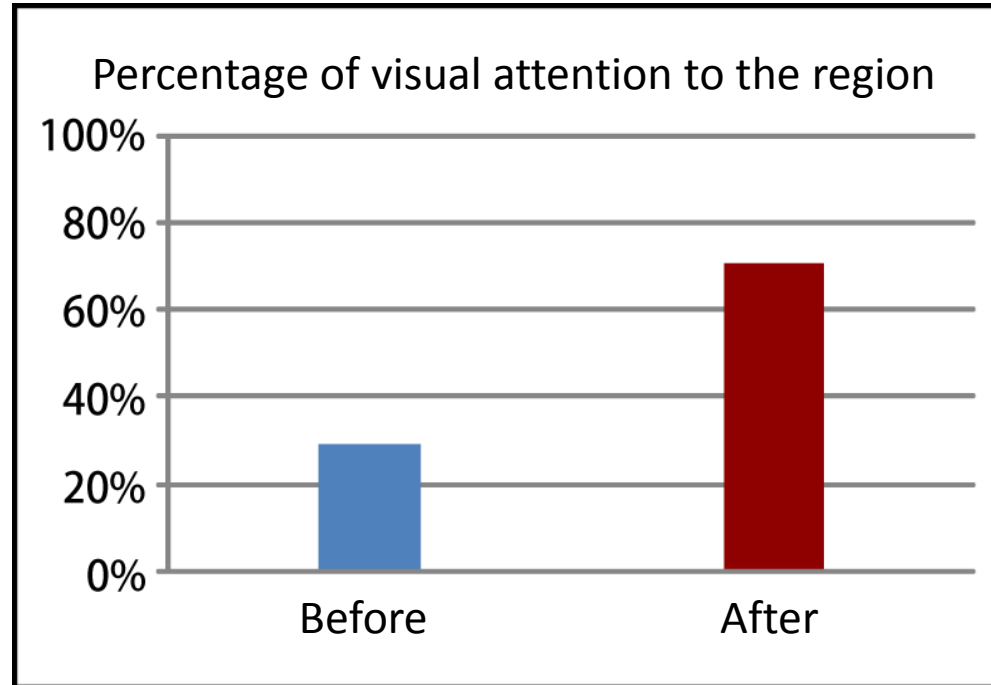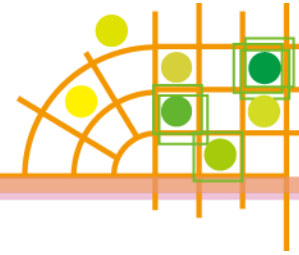
Saliency map

Evaluation using subjects

Experimental setup

- Present an edited and unedited images for 5 seconds to 10 subjects.

- If the gaze point does not move more than 0.9 degrees for 0.1 seconds, we regarded that the subject was fixating that point.
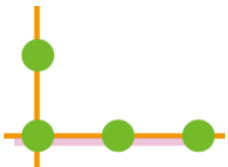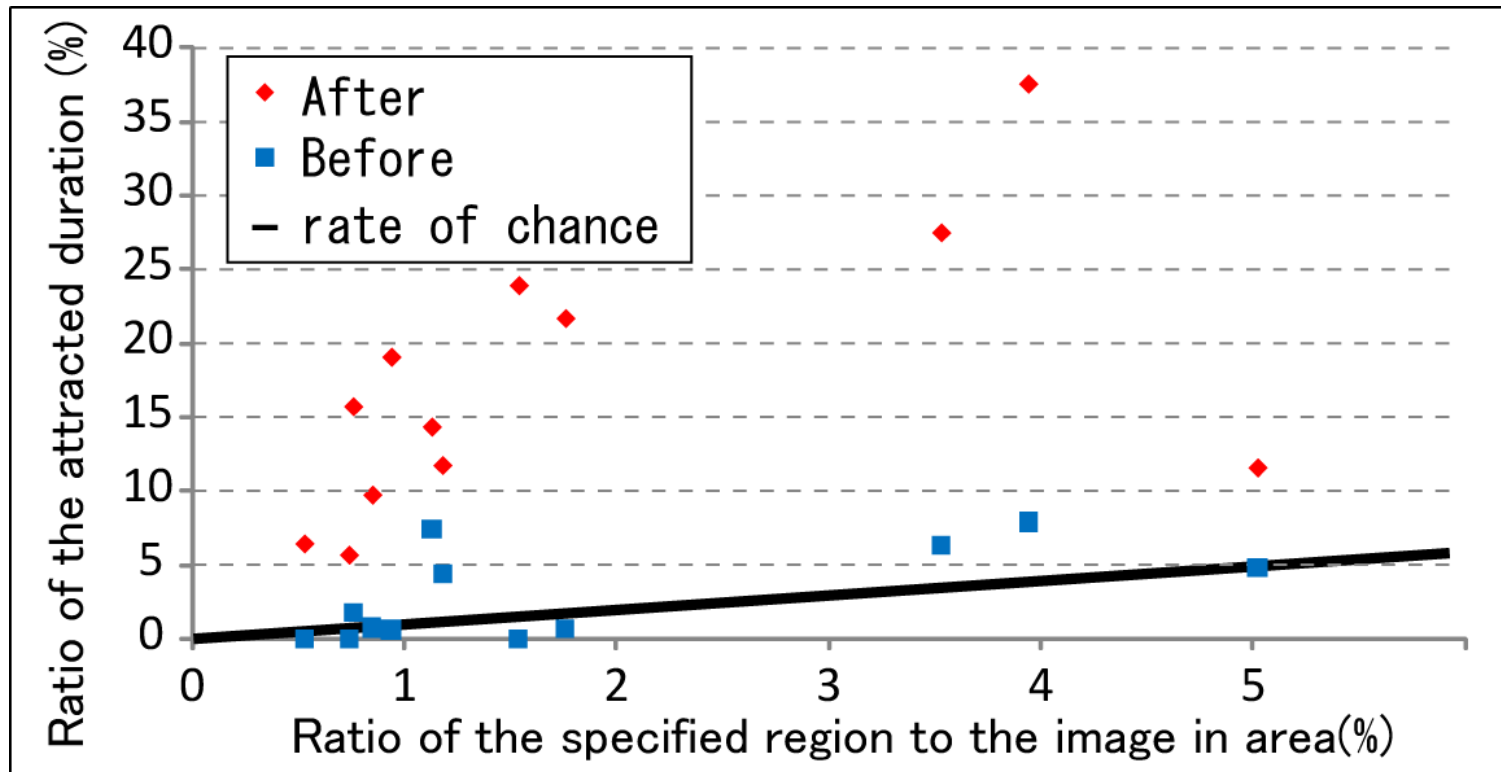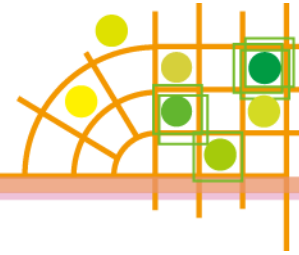
Evaluation results
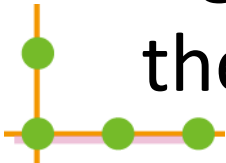

Percentage of visual attention to the region

- Average rates of drawing visual attention to a given region increase by our editing.
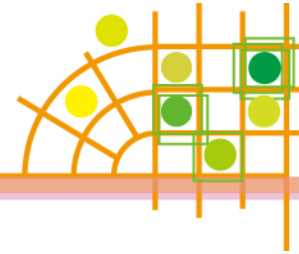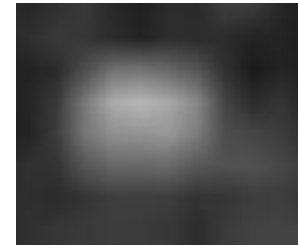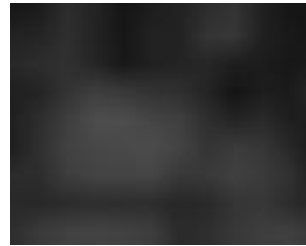
Evaluation results (cont'd)

- Significantly attracted visual attention from the statistical point of view.

Conclusion

- We presented an image editing method based on saliency for naturally guiding the user's visual attention.

  → No need to present external visual stimulus.

  → Directly modulating visual saliency.

- The rates of drawing visual attention to a given region increased by our editing.