Algorithms for Automatic Image Cropping
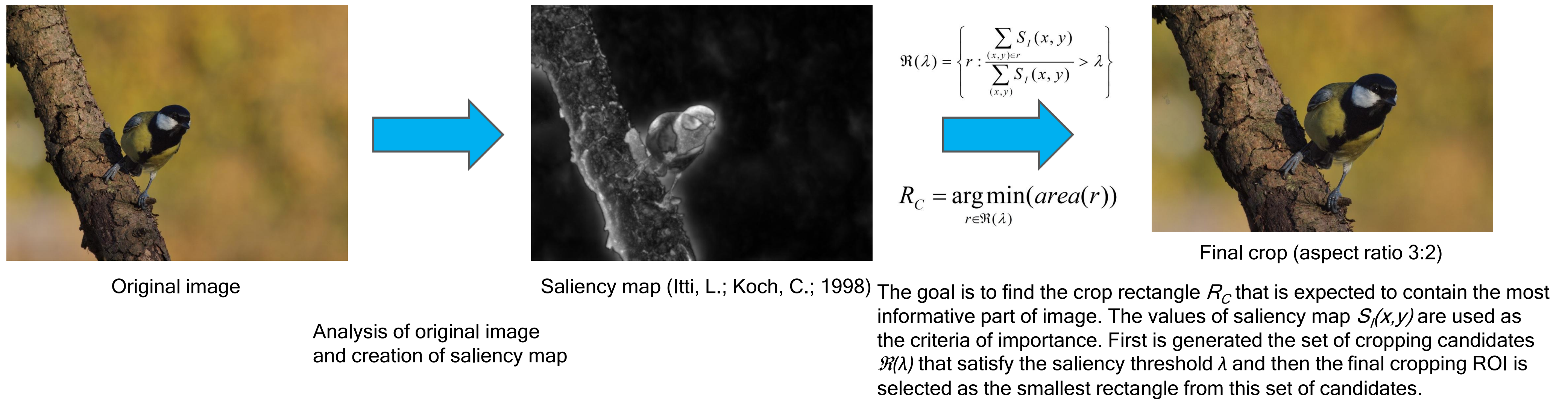
Author: Vít Ambrož, xambro15@stud.fit.vutbr.cz
Supervisor: Doc. Ing. Martin Čadík, Ph.D.
CPhoto@FIT

BRNO FACULTY
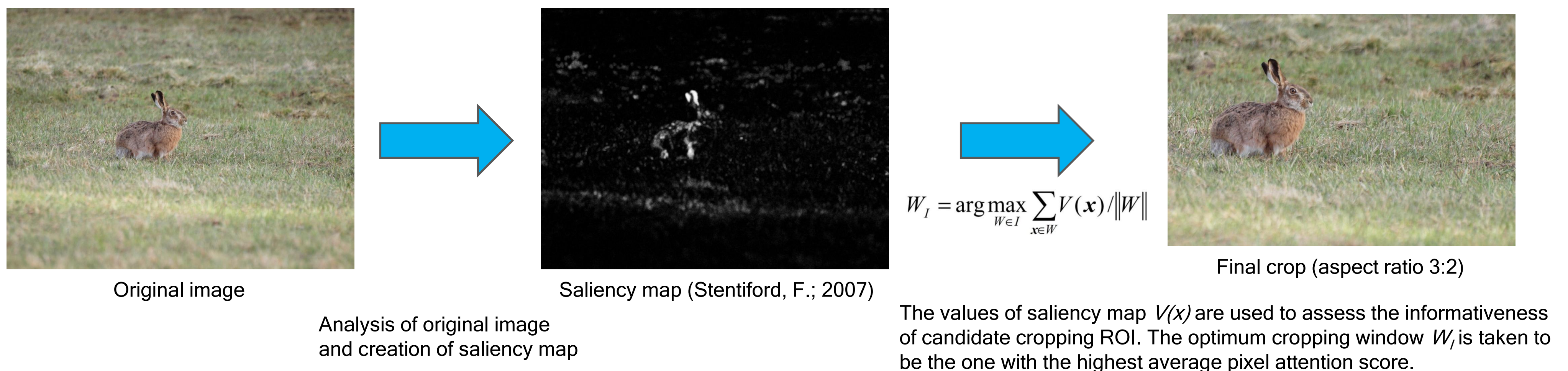UNIVERSITY OF INFORMATION
OF TECHNOLOGY TECHNOLOGY

3 existing algorithms of automatic image cropping were chosen and implemented in this work. Several cropping methods with different parameters were created for these algorithms. The algorithms of automatic image cropping could have good usability when cropping a huge number of photos or when displaying thumbnails of original photos on social sites. The implementation was realized with C++ and OpenCV.

Automatic Thumbnail Cropping and its Effectiveness (Suh et al.; 2003)

Original image

Analysis of original image
and creation of saliency map

Saliency map (Itti, L.; Koch, C.; 1998)

$$\Re(\lambda) = \left\{ r : \frac{\sum\limits_{(x,y)\in r} S_I(x,y)}{\sum\limits_{(x,y)} S_I(x,y)} > \lambda \right\}$$

$$R_C = \arg\min_{r \in \Re(\lambda)} (area(r))$$

Final crop (aspect ratio 3:2)

The goal is to find the crop rectangle R_C that is expected to contain the most informative part of image. The values of saliency map $S_I(x,y)$ are used as the criteria of importance. First is generated the set of cropping candidates $\Re(\lambda)$ that satisfy the saliency threshold λ and then the final cropping ROI is selected as the smallest rectangle from this set of candidates.

Attention Based Auto Image Cropping (Stentiford, F.; 2007)

Original image

Analysis of original image
and creation of saliency map

Saliency map (Stentiford, F.; 2007)

$$W_I = \arg\max_{W \in I} \sum_{x \in W} V(\boldsymbol{x}) / \|W\|$$

Final crop (aspect ratio 3:2)

The values of saliency map $V(x)$ are used to assess the informativeness of candidate cropping ROI. The optimum cropping window W_I is taken to be the one with the highest average pixel attention score.

Automatic Image Cropping using Visual Composition, Boundary Simplicity and Content Preservation Models (Fang et al.; 2014)

Content Preservation model is represented by saliency map of original image. The saliency map emphasizes the important objects that capture more attention of human observer. These salient objects should not be cropped out and the final crop should contain them. This model removes cropping candidates that do not achieve the minimum saliency value.

TRAINED MODEL

SVR model

Visual Composition model tries to find the cropping ROI that has the best feature of composition. This feature represents the spatial configuration of important regions and is encoded as spatial pyramid of saliency map. It is necessary to construct large dataset of well-composed images and make random crops of them to create bad-composed images. Then SVR(Support Vector regression) is used to learn to mapping from feature space to composition score. The output of SVR for a given crop candidate is the result of this model.

Original image

Saliency map
(Margolin et al.; 2013)

Final crop (aspect ratio 3:2)

$$S_{final}(C,I) = w_1 R_{compos}(C,I) + w_2 R_{boundary}(C,I)$$

Boundary Simplicity model encourages crop boundaries to pass visual simpler regions and reduces the chance of cutting through important objects. This model is represented by gradient map of original image where the edges and corners are emphasized.

First the candidates of cropping ROIs are proposed and *Content preservation model* is used to remove ROIs that miss the main salient objects. Then the final crop is selected as ROI with the best result according to combination of *Boundary Simplicity* and *Visual Composition models*. These result could be affected with definition of weights w_1, w_2 for these models to find the best crop.

Gradient map (generated using Sobel filter)