Facial Component Extraction by Cooperative Active Nets with Global Constraints

Ryuji Funayama†‡ Naokazu Yokoya† Hidehiko Iwasa† Haruo Takemura†

†Nara Institute of Science and Technology (NAIST),
8916-5 Takayama, Ikoma, Nara 630-01, Japan
‡Information Systems Product Development Laboratories, SHARP Corporation,
492 Minosho, Yamatokoriyama, Nara 639-11, Japan
E-mail: {ryuuji-f, yokoya, iwasa, takemura}@is.aist-nara.ac.jp

Abstract

This paper describes a new method for extracting facial components from a color image using cooperative active nets with global constraints. We utilize an active net model which is a region extraction method based on energy minimization principle. Each net deforms with its own energy being minimized and the position of the nets are controlled by minimizing an additional energy defined by global constraints on their placement. Our method has been experimentally shown to be robust to variations of facial size, position and rotation.

1. Introduction

Facial component extraction is very important because it has various applications such as personal identification, teleconferencing and advanced human interface. The typical methods in facial component extraction are projection-based technique and template matching[1]. They both have the advantages of algorithm simplicity, but they sometimes suffers from lighting condition, background noises and rotation of the image.

In addition to these methods, facial component extraction is achieved by using deformable templates[8][7]. It is based on energy minimization principle and requires initial templates which must be carefully set near the target facial component.

This paper describes a new face and facial component extraction method from a color image using active net models[5]. The nets deform by their own energy and move by global constraints of their placement. By cooperatively deforming models designed for facial components, it is possible to reduce extraction errors. Our method has been experimentally shown to be robust to variations of facial size, position, rotation and initial position of active nets.

2. Active Net Model

Here we briefly describe the active net models. There have been active contour models (snakes) originally proposed by M. Kass et al.[3]. Snake is defined as a contour on an image which deforms by minimizing its energy to extract an object boundary in the image. Total energy consists of internal energy, image force and external constraint forces.

Active net model by K. Sakaue and K. Yamamoto[5] is a two-dimensional extension of snakes and achieves robust region extraction using image features within the net such as textures. Energy of a net (E_{net}) is defined as:

$$\begin{aligned} E_{net} &= \int_0^1 \int_0^1 \{E_{int}(\boldsymbol{v}(p,q)) + E_{image}(\boldsymbol{v}(p,q)) \\ &+ E_{con}(\boldsymbol{v}(p,q))\} dp dq, \end{aligned}$$

where E_{int} is an internal energy, E_{image} is an image force, E_{con} is an external constraint force and $\boldsymbol{v}(p,q)$ denotes a grid point of the net. The net is organized by connecting the nearby grid points.

3. Analyzing Color Distribution of Face Region

We utilize color information to extract a face from an image. First we manually extract facial and mouth regions from face images. To characterize each region, we analyze the color distribution of the region. We use hue and saturation of HSV color space because it is desirable that features are robust to variations of brightness and color by lighting conditions and individuals[4].

By analyzing the distribution of hue and saturation in sample images of 24 persons, a two-dimensional normal distributional probability density function parameterized by values of hue and saturation is derived[2]. A value of this function represents a probability of a facial region.

Figure 1 illustrates the obtained probability density function with respect to hue and saturation, where the vertical axis represents the probability density. A similar function for the mouth region is also derived in the same manner. Figure 2 shows probability images for facial and mouth regions obtained by applying these functions on an input color image. Intensity of a pixel represents the probability that the pixel is from each region.

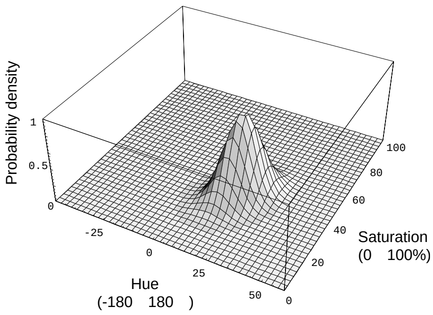
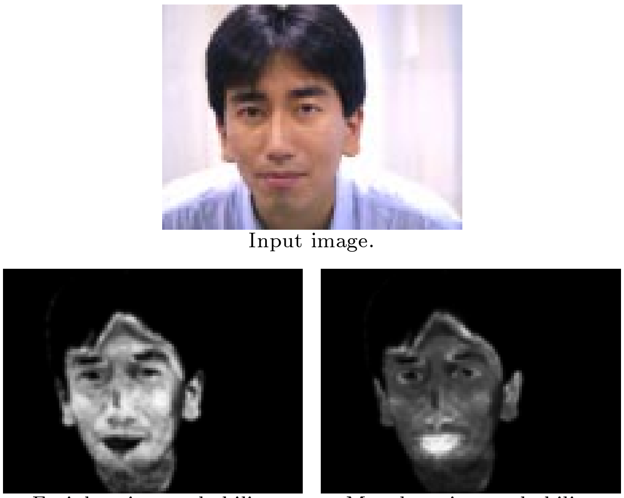


Figure 1. 2D probability density function for facial region.



Input image.

Facial region probability image.

Mouth region probability image.

Figure 2. Probability images.

4. Extraction of Face and Facial Components by Active Nets

Our algorithm is able to extract face and facial components; mouth, right and left eyes from a color image which contains a single person roughly looking front. Figure 3 schematically illustrates the flow of the algorithm for extracting face and facial components. First a net model is initialized to cover the whole image and it starts to deform for extracting a face. The result will be then used to extract mouth, right and left eyes. These nets deform in parallel and cooperatively.

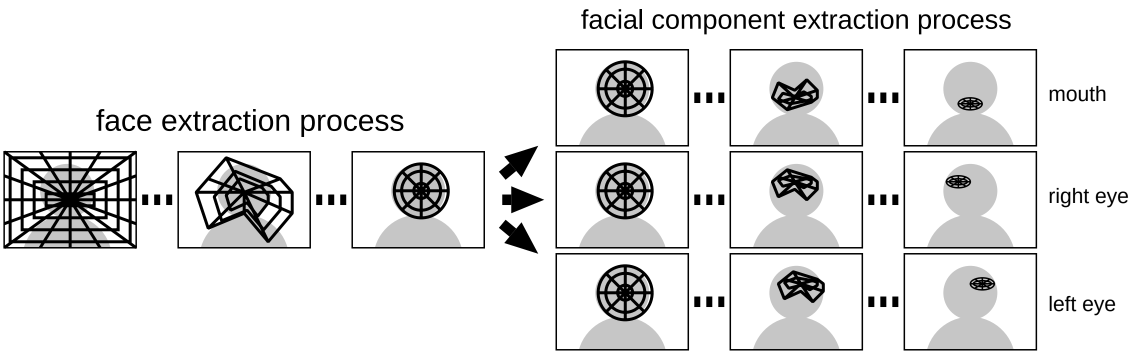


Figure 3. Illustration of the flow of extracting face and facial components.

4.1. Face Extraction by Active Net

4.1.1. Model of a Net

First we define a model of net. Our objective here is to extract a face region which is generally considered as an ellipse. Thus we define the initial net as a round shape. Figure 4 shows the model where $\boldsymbol{v}$ means a grid point of the net.

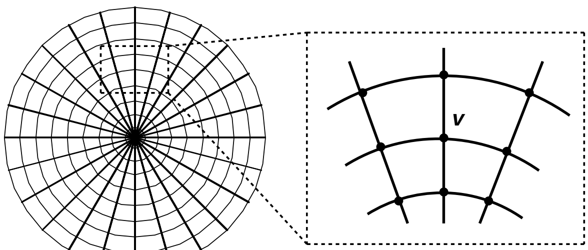


Figure 4. Model of a net.

4.1.2. Internal Energy

Internal energy for a face extraction net ($E_{face-int}$) produces a force to contract the net with its shape keeping round. This energy is explained minutely at the original paper of active net[5].

4.1.3. Image Force

Image force for the face extraction net ($E_{face-image}$) is a force to suspend the contraction of the net when it covers the facial region. It is defined by using the facial region probability image as shown in Figure 2. Denoting intensity at the grid point $\boldsymbol{v}$ of facial region probability image as $I_{face}(\boldsymbol{v})$, the energy for the image force is defined as:

$$E_{face-image} = \begin{cases} -\gamma I_{face}(\boldsymbol{v}); \\ \qquad \text{if } \boldsymbol{v} \text{ is an inner grid point,} \\ \gamma I_{face}(\boldsymbol{v}); \\ \qquad \text{if } \boldsymbol{v} \text{ is an outermost grid point,} \end{cases}$$

where γ is a positive weight.

4.2. Facial Component Extraction by Active Nets

4.2.1. Internal Energy

Internal energy here also provides a force to contract a net with its shape keeping round. In this paper, facial components mean mouth, right and left eyes. Internal energy for extracting each component ($E_{mouth-int}$, E_{re-int} and E_{le-int}) is defined as same as the energy of face extraction net.

4.2.2. Image Forces

Image forces make the nets attracted to salient features of corresponding facial component. Image forces of each region extraction net ($E_{mouth-image}$, $E_{re-image}$, $E_{le-image}$) are also defined using probability images. For instance, image force of mouth region extraction net is defined as:

$$E_{mouth-image} = \begin{cases} -\delta\Phi(I_{mouth}(\boldsymbol{v}) - I_{face}(\boldsymbol{v})); \\ \qquad \text{if } \boldsymbol{v} \text{ is an inner grid point,} \\ -\varepsilon I_{face}(\boldsymbol{v}) - \zeta\nabla|I(\boldsymbol{v})|; \\ \qquad \text{if } \boldsymbol{v} \text{ is an outermost grid point,} \end{cases}$$

$$\text{where } \Phi(x) = \begin{cases} x; & x > 0, \\ 0; & x \leq 0, \end{cases}$$

where $I_{mouth}(\boldsymbol{v})$ denotes the intensity at $\boldsymbol{v}$ in the mouth region probability image shown in Figure 2. $\nabla|I(\boldsymbol{v})|$ means edge magnitude of the input image. δ, ε and ζ are positive weights.

4.2.3. External Constraints

We define two types of external constraint forces for each facial component: E_{form} and E_{sym}. The former is a constraint force concerning the shape of facial components and the latter is a constraint for maintaining the symmetry of two eyes.

The size of facial components can be roughly estimated from the size of a face and their shapes are regarded as ellipses. Global shape constraint forces (E_{form}) are defined as forces that contract nets to the corresponding estimated ellipses (see Figure 5).

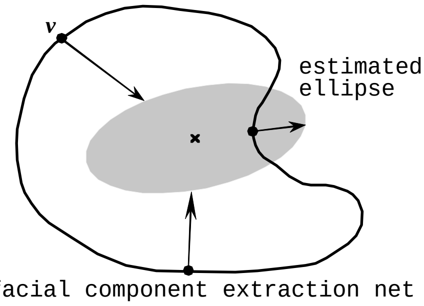


Figure 5. Global shape constraint force.

Symmetry constraint forces are defined for nets for extracting right and left eyes. They are obtained from the symmetry of both eyes. These forces make the shape of nets for right and left eyes become symmetrical (see Figure 6).

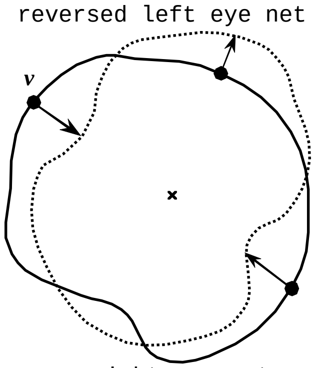


Figure 6. Symmetry constraint force.

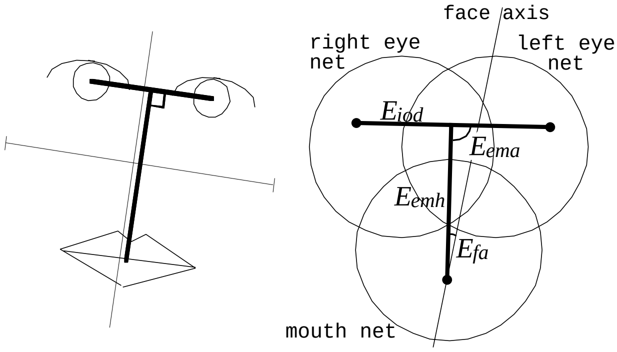


Figure 7. Placement of facial components and corresponding energy terms.

4.3. Total Energy for Extracting Face and Facial Components

The total energies for extracting face and facial components are now obtained as follows.

$$E_{face-net} = E_{face-int} + E_{face-image}.$$

$$E_{mouth-net} = E_{mouth-int} + E_{mouth-image} + E_{mouth-form}.$$

$$E_{re-net} = E_{re-int} + E_{re-image} + E_{re-form} + E_{re-sym}.$$

$$E_{le-net} = E_{le-int} + E_{le-image} + E_{le-form} + E_{le-sym}.$$

4.4. Cooperative Active Nets with Global Constraints

The nets for extracting facial components act cooperatively according to global constraints about the placement of facial components. We introduce a new energy term $E_{balance}$ defined as:

$$E_{balance} = E_{mg} + E_{reg} + E_{leg} + E_{iod} + E_{emh} + E_{ema} + E_{fa}.$$

E_{mg}, E_{reg} and E_{leg}: Forces for making the nets for mouth, right and left eyes move toward the target object regions.

E_{iod}: Force for making the distance between both eyes a specific value estimated from the width of the face (see Figure 7 and so forth).

E_{emh}: Force for making the distance between the center of both eyes and the mouth a specific value estimated from the width of the face.

E_{ema}: Force for making the line connecting both eyes perpendicular to the line connecting the center of both eyes and the mouth.

E_{fa}: Force for making the axis of face parallel to the line connecting the center of both eyes and the mouth.

The nets for extracting face and facial components are deformed by minimizing their own energy, and the positions of the nets are controlled as their placements being balanced.

5. Experimental Results

We have experimented with color images which consist of a single person in a room. Figures 8 and 9 show sequences of extracting face and three facial components, where the input image is given in Figure 2. It can be observed that the initial nets covering the whole image are correctly converging to the face, mouth, right and left eyes. In order to demonstrate the robustness of the proposed method to rotation, Figure 10 shows the results for the rotated image.

These experiments show that the proposed algorithm is effective to facial component extraction and is robust to the rotation of face.

6. Conclusion

We have proposed a new facial component extraction method using cooperative active nets with global constraints. It has been experimentally proven that:

- Facial components are extracted robustly against variations in pose and location of a face.
- Our cooperative active net models are not required to be carefully initialized.

Our method is effective if only one person is in an input image. We will extend the method to multiple face extraction by employing splitting active net models. Further investigations are also required about the determination of weights of energy terms which are at present determined empirically.

References

[1] R. Brunelli and T. Poggio. Face recognition: Features versus templates. *IEEE Trans. PAMI*, 15(10):1042–1052, 1993.

[2] R. Funayama. Face and facial component extraction by cooperative active nets. *Master's Thesis of NAIST*, NAIST-IS-MT9451096, 1996. (in Japanese).

[3] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active countour models. *Proc. ICCV*, pages 259–268, 1987.

[4] S. Matsuhashi, K. Fujimoto, O. Nakamura, and T. Minami. A proposal of the modified hsv colour system suitable for human face extraction. *The Journal of the Institute of Television Engineers of Japan*, 49(6):787–797, 1995. (in Japanese).

[5] K. Sakaue and K. Yamamoto. Active net model and its application to region extraction. *The Journal of the Institute of Television Engineers of Japan*, 45(10):1155–1163, 1991. (in Japanese).

[6] M. A. Turk and A. P. Pentland. Face recognition using eigenfaces. *Proc. CVPR'91*, pages 586–591, 1991.

[7] X. Xie, R. Sudhakar, and H. Zhuang. On improving eye feature extraction using deformable templates. *Pattern Recognition*, 27(6):791–799, 1994.

[8] A. Yuille, D. Cohen, and P. Hallinan. Feature extraction from faces using deformable templates. *Proc. CVPR'89*, pages 104–109, 1989.

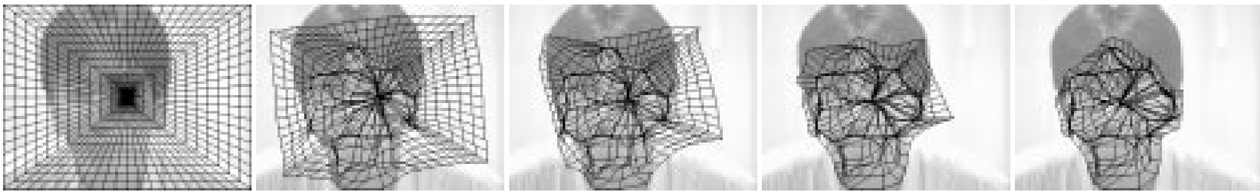


Figure 8. A sequence of active nets converging to a face region

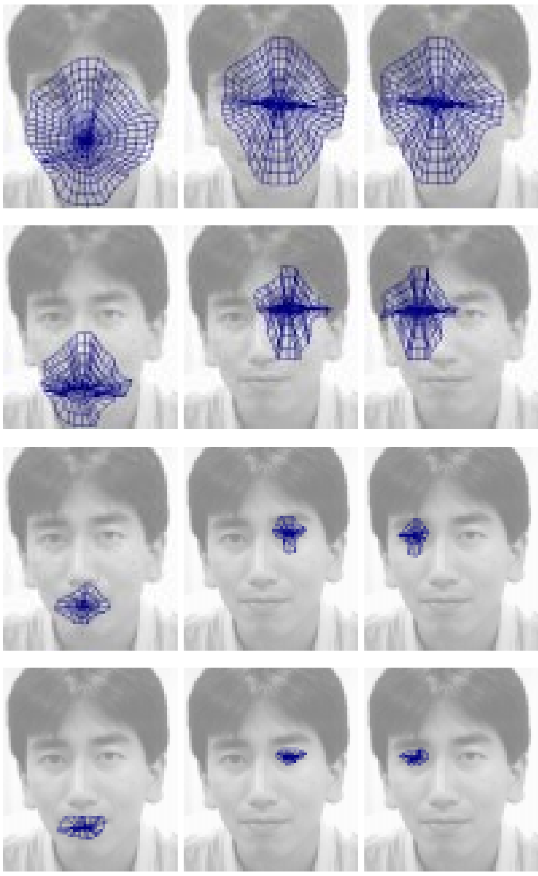


Figure 9. Results of extracting three facial components.

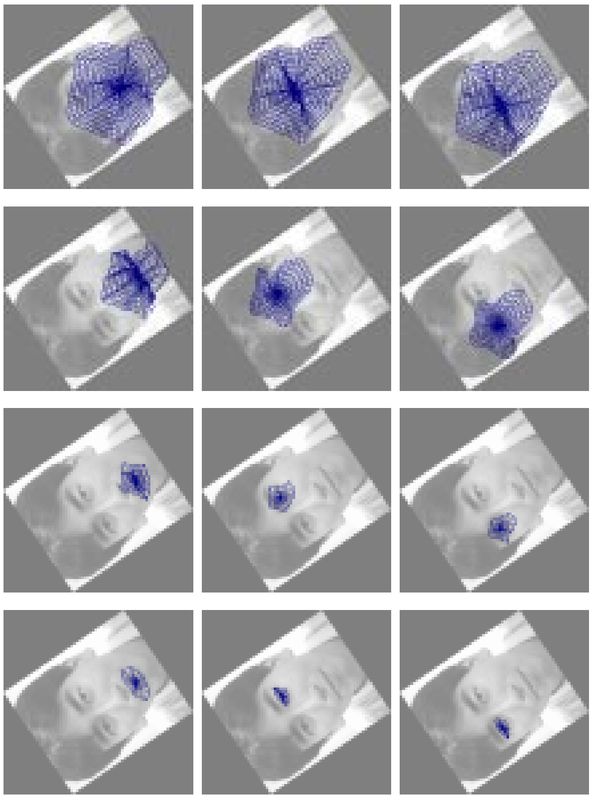


Figure 10. Results for a rotated image.