MVA'94 IAPR Workshop on Machine Vision Applications Dec. 13-15, 1994, Kawasaki

Adaptive Energy Function for Active Net

Kazuyosi YOSHINO, Mio MAKI, Toshio KAWASHIMA, and Yoshinao AOKI

Department of Information Engineering, Hokkaido University
Kita-13, Nishi-8, Kita-ku, Sapporo 060 JAPAN

ABSTRACT

This paper proposes an algorithm to extract a target region from a color image. The algorithm utilizes the active net for stable target extraction. In order to extend the original active net for color image analysis, the image energy function of the net is defined by the function, called the Ratio Histogram, of the histogram backprojection method. Since ratio histogram is computed from both color histograms of a target and an input image, the adaptive image energy function for the input image can be created. Then, since our proposed method uses target colors as the index of the target, the net can extract a target regardless of the target size and of its pose. Additionally, we refine the proposed image energy dynamically during iterations of the net to improve the net convergency. Experimental results for color images taken by a video camera show the validity of our approaches.

1 Introduction

Computer vision research has concentrated on the problem of image interpretation to supply machines with the same mechanism as the human eyes. Image interpretation is required to find an object in an image and to identify the object simultaneously with the object features, shape or colors. In the real world, solving the problem of image interpretation becomes difficult since many objects may exist in one image.

Ballard [1] emphasizes that the methods of computer vision must be simplified by dividing the image interpretation into two tasks; *location* and *identification*. The location task locates an interesting object by the object features. The identification task identifies the object of which the location is known by the object shape and by its features.

The first approach has been studied to detect image configuration, such as image segmentation. Image segmentation detects an interesting object by the object features: edges, textures, colors and so on. These methods cannot detect a object shape when noises are in an input image or when the object edges are discontinuity. Recently, the approaches based on a deformable model have been studied to detect an interesting object shape. Kass[3] introduced *SNAKES* which is a typical example of this approach. The principal purpose of SNAKES is to stably and accurately detect an interesting object shape. Consequently, SNAKES is situated in the latter computer vision task.

Active net [4], introduced by Sakaue et al, is also one of the approaches which uses a deformable model. The active net that extends the snakes model into a 2-D process detects an object using the regional features of the object besides the edge features that SNAKES uses. In the active net algorithm, if the image energy function of the net is defined with the index of an interesting object, the active net can simultaneously actualize both vision tasks. Consequently, the problem indicated by Ballard can be solved using the active net.

The original active net for a gray image often fails to extract a target region since the target index is limited. The problem is solved by applying the active net to a color image; for example, when a target has a single color, the net can extract the target region in the color image. However, most objects in the real world have some pattern; i.e. the objects are composed of multiple colors. The object colors may also be included in a background area. In this case, the net cannot stably extract the object region since the background area which includes the object colors is regarded as a part of the object.

For color image analysis, Swain and Ballard proposed *Histogram Backprojection* [2] which detected the features of an interesting object (target). Histogram backprojection creates a color histogram of the target and an input image, computes the ratio of corresponding both histogram bins, and detects the target features. The target features are the colors which have a large ratio value. The output image of the histogram backprojection is created by replacing the ratio value of each color to the pixel which had the same color in the input image. Since the colors which are regarded as the target features may be a part of the target colors, the pixels which have one of the colors are scattered on the output image. Consequently, histogram backprojection can detect the target location by filtering the output image, as shown in [2], but cannot detect the target shape.

This paper proposes a method to detect a target location and its shape from a color image using the advantages of both the algorithms, active net and histogram backprojection, together. In the method, the incorporation of the both algorithms is performed by defining the function which computed the ratio in histogram backprojection as the image energy of active net. Additionally, this paper describes a method for improving the net convergency. The method redefines the proposed image energy during iterations of the net, dynamically.

2 Region Extraction from Color Image

2.1 The Algorithm of Active Net Model

Active net is a deformable model which utilizes the network analogy of a physical region. The algorithm detects a target region by minimizing the energy function defined for the sample points, $\mathbf{v}(p,q) = (x(p,q), y(p,q))$, of the net model. In the algorithm, the net can stably extract a target region since the net has sample points also inside the net model.

The energy function consists of the internal strain energy of the net, the image energy of the target, and the external constraint energy. The internal strain energy controls the smoothness of the net structure. The image energy corresponds to the force which attracts the net to the target region and is defined from a target feature. The external constraint energy gives rise to the repulsion force. In our experiments, the external constraint energy is ignored. The total energy of the net model is computed from a linear combination of the several energy functions

$$E_{net} = \int_0^1 \int_0^1 (E_{int}(\mathbf{v}(p,q)) + E_{image}(\mathbf{v}(p,q)) + E_{con}(\mathbf{v}(p,q)))dpdq. \quad (1)$$

The active net can extract a target region, if the feature of a target is defined as the image energy. In the original research, the algorithm is applied for extracting a target region from a gray-image. The image energy is defined from the image intensity. In a gray-image, however, the net cannot often extract a target region, because color information and the other color information may be transformed to the same gray-level when a color image is transformed into a gray image. In order to solve this problem, we attempt to apply the active net to a color image and to use the colors of a target as the target index. Ordinarily, it is difficult to use the target colors for indexing the target, since an object has multiple colors and the colors are often included in a background area. In this paper, to extract a target region from a color image, we define the image energy of the active net by the ratio histogram of the histogram backprojection method.

2.2 The Algorithm of Histogram Backprojection

Histogram Backprojection is a method to detect a target position from a color image using the index of the pixel in the image. The index represents whether the pixel in the image is a part of the target region. In the algorithm, the model image of a target is registered first. The model image is the instance image in which contained a single target. Given a model image and an input image, the color histogram is obtained by counting the number of the pixels whose pixel value fall into the same bin. Using two color histograms, the index of the pixel in the input image is computed from the function called the *Ratio Histogram*. The ratio histogram R is defined as

$$R_i = Min(\frac{M_i}{I_i}, 1) \quad (2)$$

where M_i and I_i are the histogram of a model and an input image for bin i, respectively. The function returns higher values when the color of an image pixel is specific to the model image. With the function, the pixel values of the input image are replaced by the value of R for their colors. Figure 1 (c) shows the R image of the input image (figure 1 (a)) for the model (figure 1 (b)). In figure 1 (c), dark pixels have a higher value of the function R.

As shown in figure 1 (c), pixels which have a high value of R are scattered in the output image. Swain has detected the target location by filtering the output image with a circular mask.

2.3 Active Net Utilized Histogram Backprojection

In order to apply the active net to a color image, the image energy of the active net is defined by the ratio histogram of the histogram backprojection method and is written as

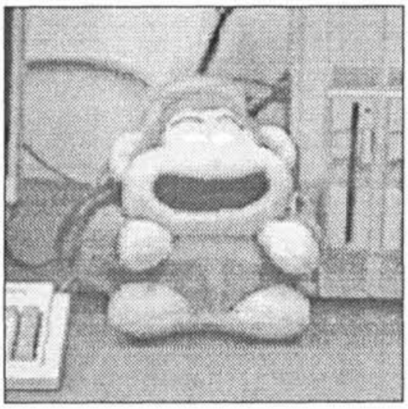
$$E_{image} = \omega R_{c(x,y)} \quad (3)$$

where $c(x, y)$ represents the color (RGB value) in an input image at a lattice point position (x, y) of the net and ω is a coefficient. If $\omega < 0$, the net will be attracted to the region which has a higher value of R in an output image.

3 Experimental Results

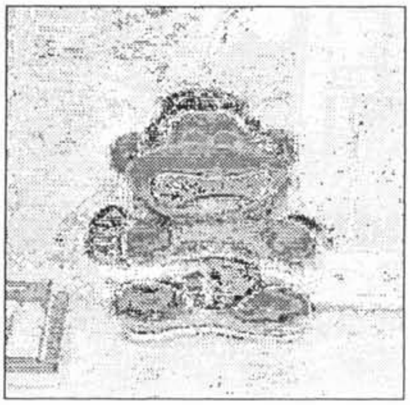
This section shows two experimental results using the active net which introduced the ratio histogram of the histogram backprojection method as the image energy of the active net method. In our experiments, the color images were taken by a video camera and were digitized at RBG of 24-bit. In the first experiment, the size of the input image and the model image are 256×256 and 55×65, respectively. In the second, the size of the input and model image are 300×300 and 65×95, respectively. The initial position of the active net model was set as covering the whole input image. In the figure of the experimental results, the active net model is illustrated by a black line.

In the first experiment, the target location and its region are detected from an input image. The target is a *Gorilla* mascot. The input image is such that the characterization of the target is difficult since the color of the target and that of the background



(a)

(b)



(c)

Figure 1: A result of Histogram Backprojection. (a) and (b) show an input image and a model image, respectively. (c) shows the R image of (a) for (b).

are transformed to the same gray level in the gray image, although their colors are different in the color image. Figure 2 shows the result of the target extraction using the proposed active net algorithm. In figure 2, the input image and the model image are as shown in figure 1 (a) and (b), respectively. In the experiment, since the target colors are not in the background area, these colors are useful for the target feature. The experimental result is satisfactory.

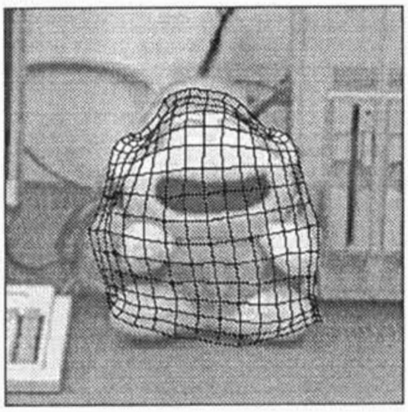
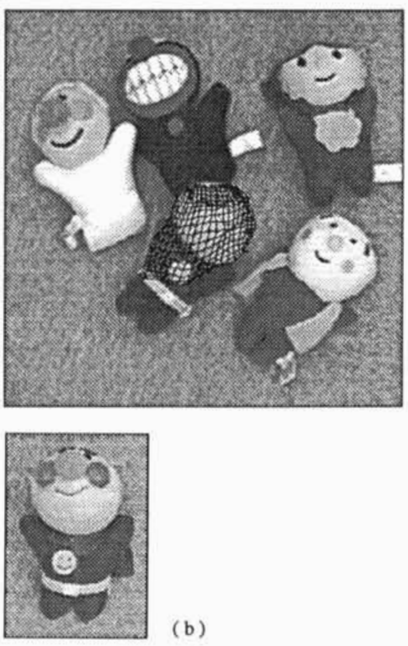


Figure 2: A result of target extraction using the proposed active net. The input image and the model image are shown in figure 1 (a) and (b), respectively. The active net model is illustrated by the dark line.



(a)

(b)

Figure 3: A result of target extraction. (b) shows the final shape of the proposed active net model (dark line) which extracted the target (a) in the input image (b).

In the second experiment, the proposed active net is applied to an input image which includes a target and other objects which have the same color as the target. Figure 3 (b) shows the experimental result for the model image (a). In the case where a target color is in the background area, since the color is excluded from the target feature, i.e. the color has a low value of the image energy, the net is not influenced by the background color and can extract only the target region. However, the target shape extracted by the proposed active net is a little different from the real target shape in the input image.

4 Improvement of Active Net Convergency

4.1 A Problem of The Proposed Active Net

Figure 4 shows the R image of figure 3 (a) for figure 3 (b). As the input image is used in the second experiment above, when a target color is included in the background area, the value of the image energy for the color is low. Consequently, our proposed active net can extract the target region without being influenced by the background color. As shown in figure 4, however, since the value of the image energy for the target color also is low, the net cannot extract the accurate target region.

4.2 Elimination of The Background Influence

In order to solve the problem, the values of the image energy for a target color need to be increased. The following two methods are considered from the equation (2):

1. a method which expands the size of a model image to increase the value of M_i in the equation (2), or
2. a method which reduces the size of an input image to decrease the value of I_i.

In the first method, since the numerator value in the equation (2) increase by expanding the model image size, the image energy value in the target region increase; however, the image energy value in the background area, which includes the same color as in the target, also increase. Consequently, This paper uses the second method, because the target region extraction becomes more difficult using the first method.

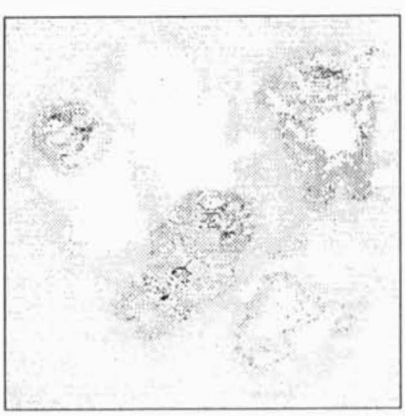


Figure 4: The R image of figure 3 (b) for figure 3 (a).

The second method decreases an input image size by excluding the background area from the input image. In the method, if the target position in an input image is known, it is easy to exclude the background area in the image. However, since our purpose is to extract the target position and its region, it is impossible to exclude the background area in advance. To solve the problem, we introduce an algorithm which excludes a background area based on the behavior of the active net.

4.3 Adaptive Image Energy Function

In the active net algorithm, the inside net points are attracted in a target region by the image energy; the contour points wrap the target region; then, the net converges. Consequently, the target region exists inside the net model during iteration for the active net. In this paper, we propose an algorithm which redefines the image energy of the active net dynamically, from the net shape during iteration; i.e. the image energy is altered by regarding the image which is included in the net shape as the input image.

Figure 5 shows the R image of figure 3 which is computed from the net shape at 1000 iterations. As shown in figure 5, the target shape appears clearly, since the R values on the target region increase by our algorithm above. The R values on a part of the background area also increase; however, since this part is already outside the net, the net is not influenced by the part.

Figure 6 shows the final shape of the proposed active net applied to figure 3 (a). As shown in figure 6, our proposed active net can extract an accurate target region by altering the image energy dynamically.

5 Conclusion

This paper has described a method to detect a target location and its shape from color image. Their detection is performed using active net introduced the image energy which is defined by the ratio histogram of histogram backprojection. Additionally, in order to improve the net convergency, this paper has shown a method to dynamically redefine the proposed image energy during iterations of the net. Since our proposed method uses target colors as the index of the target, the target location and its shape can be detected independently of the target size and of its pose in an input image; however, if target colors in a model image and in an input image are differed by illumination, target detection becomes difficult.

Further work will consider determining the optimal size of a model image and solving the problem of color constancy.

References

[1] Dana H. Ballard : "Animate Vision", Artificial Intelligence, Vol.48, pp.57-86 (1991)

[2] Michael J. Swain and Dana H. ballard : "Color Indexing", International Journal of Computer Vision, Vol.7, No.1, pp.11-32 (1991)

[3] Michale Kass, Andrew Witkin and Demetri Terzopoulos : "Snakes : Active Contour Models", International Journal of Computer Vision, Vol.1, No.4, pp.321-331 (1988)

[4] Katsuhiko Sakaue and Kazuhiko Yamamoto : "Active Net Model and Its Application to Region Extraction (Japanese) ", JITE of Japan, Vol.45, No.10, pp.1155-1163 (1991)

Figure 5: The R image of figure 3 (b) for figure 3 (a) at 1000 iterations of the active net.

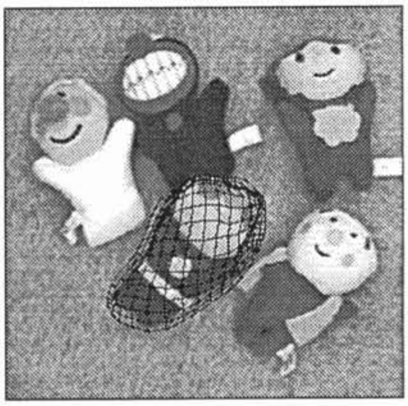


Figure 6: The final shape of our proposed active net applied to figure 3 (b).