Shape Modeling from Multiple View Images Using GAs

Satoshi KIRIHARA and Hideo SAITO

Department of Electrical Engineering, Keio University
3-14-1 Hiyoshi Kouhoku-ku Yokohama 223, Japan
TEL +81-45-563-1141 (ext.3310), FAX +81-45-563-2773
E-mail: kiri@ozawa.elec.keio.ac.jp, saito@ozawa.elec.keio.ac.jp

Abstract

Shape modeling is a very important issue for many study, for example, object recognition for robot vision, virtual environment construction, and so on. In this paper, a new method of object modeling from multiple view images using genetic algorithms (GAs) is proposed. In this method, a similarity between model and every input image is calculated, and then the model which has the maximum similarity is found. For finding the model of maximum similarity, genetic algorithms are used as the optimization method. In the genetic algorithm, the sharing scheme is employed for efficient detection of multiple solution, because some shape may be represented by multiple shape models. Some results of modeling experiments from real multiple images demonstrate that the proposed method can robustly generate model by using the GA.

1 Introduction

Many methods have recently been studied for 3-D object modeling. These methods can be applied to object recognition for robot vision [1], construction of virtual world [2][3], and so on. Especially, virtual world is significant subject for communication using computer networks. Many previous modeling method use range data of the object shape which can generally be obtained by range finders [4] [5].However, range finders are not popular tool, and they can not obtain the surface texture data which can be used for synthesizing images presented in virtual space (image-based rendering).

Therefore, many methods for recovering range information from 2-D images taken with CCD camera have been studied [6][7]. Generally, because these methods use some multiple images, some corresponding points between each image must be detected. If the corresponding points are detected exactly, the range information can be recovered by triangulation. However, the detection of corresponding points is known as a difficult problem and thus have been studied hardly. It is especially difficult to make correspondence between multi-view images because of the occlusion and the inconsistency of background scene.

If we consider that the object problem is not estimating accurate range information but generating accurate object models, we don't have to recover the range data. For example, we can obtain the object model by the use of interactive operation system [8] in which generation of hypothesis of models and verification of the hypothesis on the multiple view images are repeated interactively. Such kind of modeling system does not require the range data of the object but verification of the generated hypothesis of models by the human.

In this study, we intend to make automatic modeling system based on the repeated operation of generation and evaluation of hypothesis. For searching the best hypothesis of model efficiently, we employ genetic algorithms (GAs) [9]. In our method, the model matching to every input images are found by applying GAs which repeat evaluation

of hypotheses of the models. For the evaluation, similarity between the model and the input image at each view point is calculated, and then the model having the maximum evaluation is found by GAs.

In our previous article [10], the concept and the early results are presented, but the quality of the results are not sufficient. In this paper, we have made significant improvement for the algorithms and obtained high quality results, which are shown in the following parts of this paper.

2 Proposed method

In this study, we assumed that the object is the polyhedral such as an artificial building. Under this assumption, the object can be expressed by multiple model shapes of triangular prisms and rectangular prisms. Input images are taken by CCD camera from multiple view. The problem is estimation of shape, position, and pose of the models from the input images. **Fig. 1** shows the scheme of the shape modeling assumed in this paper.

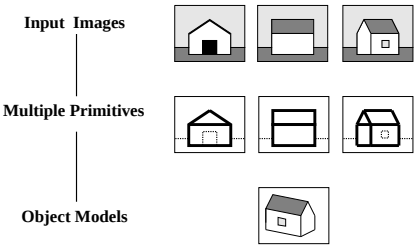


Fig. 1. Scheme of the shape modeling.

In this method, modeling is performed by maximizing an evaluation function using GAs. The evaluation function is described later.

2.1 Model parameters

As described above, we employ the model shapes of triangular prisms and rectangular prisms for representing the object surface.

We define the model parameters as the following:

1. model category (rectangles or triangles) (A),
2. length to x, y, z axis (lx, ly, lz),
3. rotation angle around y axis (bt),
4. position of center of gravity (xc, yc, zc).

2.2 Modeling by GAs

Fig. 2 shows the flow of the modeling process by GAs. First, a group of strings is produced at random. These strings represent hypothesizes of the model. All strings are evaluated, then the genetic operations are applied to them. The string which has the best evaluation is searched by repeating the genetic operation. The model parameters represented by the best strings is regarded as the final solution.

Definition of string For using GAs, the object model must be expressed as a string. As described in the previous section, the object model can be represented by some simple parameters in this method. Each parameters has 8 bits binary except for the parameter of model category. The parameter of model category has only 1 bit, because it is used for only distinguishing between rectangles and triangles. Then, the strings

consist of 57 bits binary. For corresponding the hamming distance of the string to the distance of the parameter value,

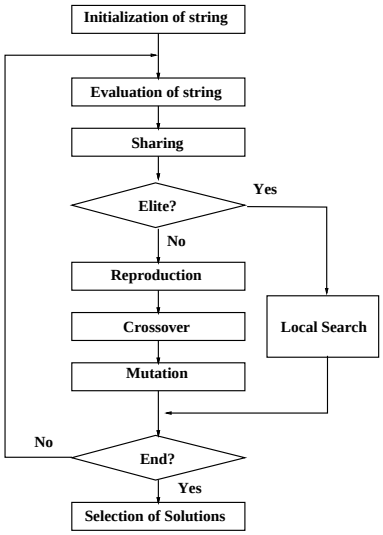


Fig. 2. The flow of the modeling process by GAs

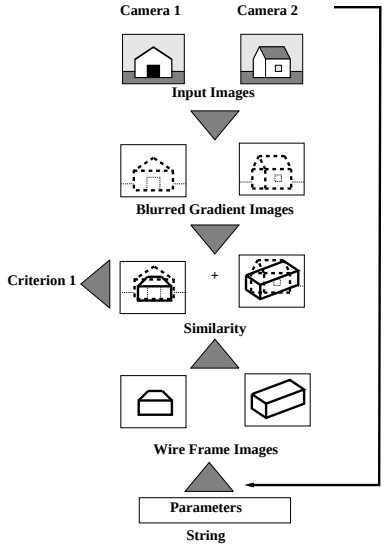


Fig. 3. Similarity criterion

Evaluation function We define evaluation function which represents similarity between input images and the model hypothesis. The following two criterion are used in this method.

1. Similarity between the input multiple view images and the synthetic wire frame images as shown in **Fig. 3** (f_1).
2. Consistency of texture patterns on the model plane which are given by multi-view images as shown in **Fig. 4** (f_2).

First criterion (f_1)is defined as the following equation.

$$f_1 = \sum_i \frac{\sum_{x,y} B_i(x,y) S_i(x,y)}{\sqrt{\sum_{x,y} B_i^2(x,y)} \sqrt{\sum_{x,y} S_i^2(x,y)}} \tag{1}$$

where S_i represents wire frame image of the model at ith view, and $B_i(x,y)$ is blurred gradient image of the input image at ith view. $B_i(x,y)$ is calculated as

$$B_i(x,y) = G(x,y) * \sqrt{\{\frac{\partial I_i(x,y)}{\partial x}\}^2 + \{\frac{\partial I_i(x,y)}{\partial y}\}^2}, \tag{2}$$

where G(x,y) is Gaussian operator and $I_i(x,y)$ is input images at ith view.

The second criterion (f_2) is defined as the following equation.

$$f_2 = -\sum_{x,y} \sum_i \{T_i(x,y) - \frac{\sum_j T_j(x,y)}{p}\}^2 \tag{3}$$

where $T_i(x, y)$ is represents texture pattern back-projected from ith input image, and p is the number of the images back-projected on the model plane. If the model parameters are exact, every back-projected texture must be the same, then f_2 have to be 0 (maximum value).

As these two criterion, the total evaluation function f is

$$f = f_1 + \alpha f_1 f_2, \tag{4}$$

where α is a weighting constant.
The second criterion (f_2) supports the first criterion (f_1). We previously used the following equation

$$f = f_1 + \alpha f_2, \tag{5}$$

which sometimes gives high evaluation f even if the model does not correspond to the object input images, because f_2 can have the high evaluation value when the object surface and the background have the smooth texture. Therefore, the second criterion (f_2) support the first criterion (f_1), then the second term ($\alpha f_1 f_2$) cannot have high evaluation until f_1 has high evaluation to some degree.

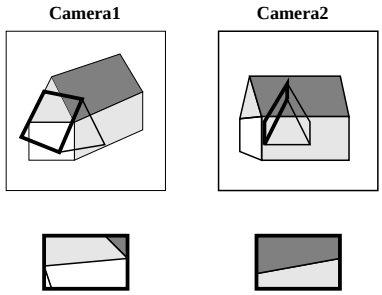


Fig. 4. Consistency of texture in criterion 2

Sharing Although GAs are generally used for finding an single solution having maximum evaluation, they can be used for finding multiple solutions simultaneously by some modifications of the algorithms. Sharing method [9] is one of the popular modification for finding multiple solutions efficiently. We employ the sharing method for finding multiple models which are included in the input images.

Sharing is an operation that decreases evaluation of strings if the strings are similar to other strings in the population of strings. By this operation, strings tend to have various values, which means that multiple solutions can be simultaneously found.

Original evaluation of a string is modified according to the number of the neighbor strings. The relationship between the modified evaluation and the original evaluation is shown in the following equation.

$$E_s(x_i) = \frac{E(x_i)}{\sum_{j=1}^{n} s(d(x_i, x_j))}, \tag{6}$$

$$\text{where } s(d) = max(1 - d/\sigma, 0),$$

x_i, x_j : ith and jth strings, $E(x_i)$: original evaluation of x_i
$E_s(x_i)$: modified evaluation of x_i, $d(x_i, x_j)$: distance between x_i and x_j
$s(d)$: sharing function of d, σ : constant determining effect of sharing
n : number of the strings in a population

In general, σ is a constant parameter. In our method, however, σ is changed in proportional to the number of generations in GAs. In the earlier generation, sharing operates effectively and many solutions can be held in the string population by the use of higher σ. In the later generation, σ has lower value, then, local optimization can be realized around the multiple solutions.

Genetic operations The flow of the modeling process by GAs is already shown in **Fig. 1**. First, the strings are initialized at random. The number of strings are 256.

Next, evaluation value is calculated for each string by the use of equation 4. The original evaluation is modified by sharing function as described in the previous section. Some elite strings are then selected according to modified evaluation. The selected elite strings are improved by a local search method. In this way, the strings having higher evaluation are selected as the elite strings.

Other strings which are not selected as elites are applied to genetic operations. Genetic operations are reproduction, crossover and mutation. First, the parent strings are selected according to modified evaluation. Then, the offspring strings are generated by one-point crossover. Some bits selected at random and reversed by mutation. This process is repeated.

After repeated this process for the number of generations defined in advance, some strings are selected as the object models. In this method, the number of the models in the input images is not given to the algorithm. Therefore the object models is selected by following criterion.

1. Selecting the string having the texture evaluation (f_2) under a threshold level.
2. Selecting the string having enough distance to the object models which are already selected.

3 Experiments

We performed some experiments for demonstrating that the proposed method is efficient for object modeling. In this experiments, some real images taken by CCD camera. The objects are a locker, a model of a roof of a house, and two tissue boxes.

These objects are located in natural environment, so the taken multi-view images have various background scene. Input images is 256×256, 8bit gray-scale. The camera parameters of every image are previously measured by initial experiment.

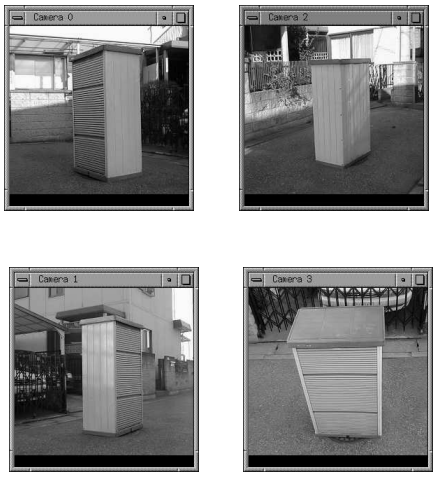
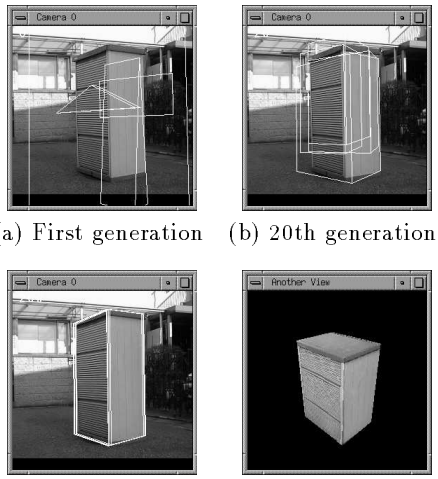


Fig. 5. Input images



(a) First generation (b) 20th generation

(c) Final generation (d) Other views

Fig. 6. Result of first experiment

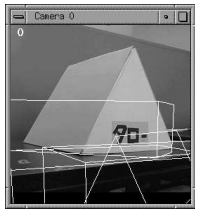
First, the experimental results for the locker are shown. **Fig. 5** shows the input multi-view images. In the first generation, candidates of solution are selected at random as shown in **Fig. 6**(a). They are getting to be close in the twenties generation as shown

in **Fig. 6**(b). In the final generation, the solutions are optimized completely as shown in **Fig. 6**(c). Then, we obtain the object model . Using the object model, we calculate the image form other view points (**Fig. 6**(d)).

As shown in the edge images (**Fig. 7**), it is very difficult to make correspondence of these edges or points between the multi-view images , because many background edges in one input image cannot exist in other input images. Since the proposed method does not require the correspondence between the multi-view images, the exact modeling result can be obtained.

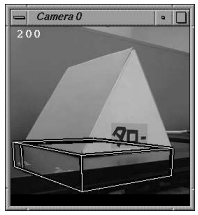
Next, we show the experimental results for a model of a roof of a house in **Fig. 8**. In this experiment, the object is expressed by the one triangular prism model. In this case, only small texture variation is included in the background in the input images, that is different from the first experiment for the locker. As described in the section of the evaluation function, if we use the simple evaluation function of eq.(5), the strings converge some fail solutions as shown in the object in **Fig. 9**(b). However,by using the evaluation function of eq.(4), the exact model can be obtained as shown in **Fig. 8**(b) although there are little texture variation.

Fig. 7. Edge image



(a) First generation (b) Final generation

Fig. 8. Result of second experiment



(a) Edge image (b) Example of fail solution

(a) First generation (b) Final generation

Fig. 10. Result of third experiment

Fig. 9. Result of simple evaluation function

Third, we show the experimental results for two tissue boxes. In this experiment, the object in the input images is needed to be represented two rectangle models. In the first generation, candidates of solution are selected at random as shown in **Fig. 10**(a). The result of modeling is shown in **Fig. 10**(b) .

The object consists of two tissues in this case, such as an artificial building is generally expressed by combination of some rectangle and triangle. The modeling results for two models have a larger modeling error than the modeling results for one model. This error is caused by the some difficulty in obtaining the multiple solution via GAs with sharing method. This will be solved by the future study.

In this study, GAs play the important role for obtaining the shape model efficiently

because of performance of optimization of GAs.

Fig. 11 shows the distribution of the evaluation function for valuable parameters. Because the distribution for all parameters can not be displayed, we show the distribution in respect to the parameters lx and ly. It is obviously recognized that the evaluation distribution has a lot of peaks. For optimizing in such evaluation distribution, the conventional methods are not suitable because they tend to give wrong solutions of local peaks. As described in papers on GA's applications, GAs are powerful algorithm in such situations.

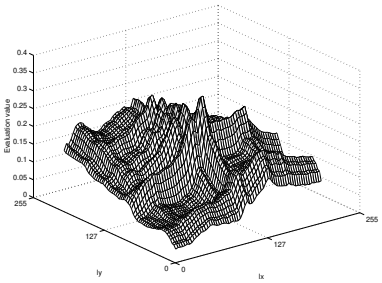


Fig. 11. The distribution of the evaluation function

4 Conclusion

In this paper, we propose a method for object modeling from multiple view images using genetic algorithms (GAs). In the proposed method, generation and verification of model hypothesis are repeated, and the hypothesis which has the best evaluation is efficiently searched by the use of GAs. Some results of object modeling experiments from real multiple images demonstrate that the proposed method can robustly generate model by using GAs.

References

1. Mark D. Wheeler and Katsushi Ikeucti : "Sensor Modeling, Probabilistic Hypothesis Generation, and Robust Localization for Object Recognition". IEEE trans. PAMI, Vol. 17, No.3, pp.252-265 (1995)
2. Carlo Tomasi and Takao Kanade : "Shape and Motion from Image Stream : a Factorization Method". Technical Report in CMU, CMU-CS-92-104 (1992)
3. Camillo J. Taylor and David J. Kriegman : "Structure and Motion from Line Segments in Multiple Image". IEEE trans. on PAMI, Vol.17, No.11, pp.1021-1032 (1995)
4. B.Parvin and G.Medioni : "B-rep from Unregistered Multiple Range Image". in Proc.of the 1992 IEEE International Conf. on Robotics and Automation, Nice, France, May, 1992, pp.1602-1607 (1992)
5. T.Masuda, K.Sakaue and N.Yokoya : "Registration and integration of multiple range images for 3-D model construction". Proceedings of the 13th ICPR, vol.1, pp.879-883 (1996)
6. S.D.Cochran, G.Medioni: "3-D Surface Description from Binocular Stereo". IEEE Trans. on PAMI, Vol.14, No.10, pp.981-994 (1992)
7. W.B.Seales, O.D.Faugeras: "Building Three-Dimensional Object Models from Image Sequence". COMPUTER VISION AND IMAGE UNDERSTANDINDING, Vol.61, No.3, pp.308-324 (1995)
8. URL: http://www.wbs.or.jp/bt/sel/skv/index.htm
9. D.Goldberg: Genetic Algorithms in search, Optimization and Machine Learning, Addison-Wesley (1989)
10. H. Saito and M. Mori: "Object Modeling from Multiple Images Using Genetic Algorithms". Proceedings of the 13th ICPR, Vol. IV, pp.669-673, (1996)