

<http://ansinet.com/itj>

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSI*net*

Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

A Reconstruction Method for Disparity Image Based on Region Segmentation and RBF Neural Network

^{1,2}Yu Shuchun, ¹Yu Xiaoyang, ¹Shan li'na, ¹Zhang Yuping, ¹Shen Yongbin, ¹Tian Miaomiao,
¹Fan Shenshen and ¹Huang Haixia

¹School of Measurement-Control Technology,
College of Measurement-Control Technology and Communications Engineering,
Harbin University of Science and Technology, Harbin 150080, China

²The Higher Educational Key Laboratory for Measuring and Control Technology and Instrumentations of
Heilongjiang Province (Harbin University of Science and Technology), Harbin 150080, China

Abstract: The Reconstruction for disparity image is the key technology for 3D image restoration in stereovision field. However, the data volume of disparity images is so large and the topological structures of disparity images is so complicated that reconstruction for disparity image is very difficult. We had done a lot of works for the sake of constructing a new method to reconstruct disparity image. In this study, a reconstruction method for disparity image based on region segmentation and isomorphic RBF(Radical Basis Function) neural network is presented. First, the disparity image is divided into some regions with adjustable threshold and edge detection. Next, reconstruction based on RBF neural network is carried out in every region, in which process the disparity point clouds are optimized. Then, all the regions are connected and the reconstruction result is obtained. When reconstruction based on RBF is executed in different regions, trainings are carried on with different resolution data according to the complexity of the structures of different regions. Experimental results show that the method proposed in this study can attain reconstruction results of high quality effectively.

Key words: Disparity image, stereo vision, RBF neural network, 3D reconstruction, point clouds

INTRODUCTION

3D information measurement and restoration technology based on stereovision has aroused extensive interest (Grant and Frank, 2010). The extent of approximation to the original scenery depends on the results of 3D reconstruction on the disparity image in the output form of stereovision. Nowadays, there are many accessible methods in 3D reconstruction, such as surface fitting reconstruction (Atkinson *et al.*, 2009), piece-wise linear reconstruction (Raissi *et al.*, 2008). Reconstruction based on physical features (Alfredo *et al.*, 2010) and reconstruction based on neural network (Brovko *et al.*, 2009). After the introduction of radical basis function to the reconstruction field, the algorithm based on RBF neural network had worked successfully with higher accuracy in reconstruction and a fast convergence speed (Guo-Hui *et al.*, 2008).

Including the adopted algorithms, 3D reconstruction for the disparity image has direct relation to its data

structure, which is also a key factor affecting the speed and quality of reconstruction. The data volume for disparity images is so large that it becomes a test for any reconstruction algorithm. However, the relevancy of the topological structures is distinct, which is reflected directly from the images. But anterior researchers commonly take the whole disparity image into account, and ignore assistant function of topological structures in reconstruction. If we can make good use of this feature, divide a disparity image into several regions properly and carry on the reconstruction process in each region, respectively, it will do great help to speed the reconstruction algorithms and improve the quality of the final results.

In this study, we take the idea of partition algorithm for reference, propose a reconstruction method for disparity image based on region segmentation and RBF neural network with the focus on RBF neural network technology.

Corresponding Author: Yu Shuchun, P.O.B. 453 No.52 XueFu Road, Nan Gang District, Harbin University of Science, Technology (West), The Higher Educational Key Laboratory for Measuring and Control Technology and Instrumentations Heilongjiang Province, Peoples Republic of China, 150080
Tel: 86-0451-86392369 Fax: 86-0451-86392308

THEORY OF RBF NEURAL NETWORK

Considering that the core part of the proposed method is performed on RBF (Radical Basis Function) neural network, we introduced the fundamental theory first.

RBF (Radical Basis Fu-ction) neural network is an important forward network. For some regional areas in the input layer, only a few weight values can affect the outputs of the network, therefore it is a local approximate network. RBF network are better than conventional BP network in terms of approximation capacity, classification and it is 10^3 or 10^4 times the learning speed of general BP network, since only few connection weights are to be adjusted in the procedure. Theoretically, it is capable of approximating continuous function and its all derivative at whatever accuracy.

Structurally, RBF network generally includes an input layer, a hidden layer and an output layer. The input data maps directly into hidden layer without weights, in which process some adjustments for the parameters of the radical basis function are required. It's non-linear. In the process that data in hidden layer maps to output layer, some modifications for the connection weights are required. It's linear. RBF network is generally shown as Fig. 1.

In the network, the output can be obtained in the following formula:

$$Y_j = w_{j0} + \sum_k w_{jk} \Phi(\|x_j - c_k\|) = \sum_{k=0}^{n_c} w_{jk} \Phi(\|c_k - x_j\|) \quad (1)$$

where, x_i ($i = 1, 2, \dots, L$) is the i th input, c_j ($j = 1, 2, \dots, n_c$) is the j th center of radical basis function, n_c is the number of neurons in hidden layer, the w_{ij} is the connection weight between hidden neuron j and output neuron k , the w_{i0} is the i th threshold of target neuron, the Y_{ij} is the output of the j th neuron to the input of the i th neuron.

Here, Φ is radical basis function, it may be several basis functions as follows:

- Gaussian function $\Phi(r) = e^{-r^2/2\sigma^2}$
- Thin-plate spline function $\Phi(r) = r^2 \lg(r)$
- Multi-quadric function $\Phi(r) = (r^2 + c^2)^{1/2}$
- Inverse multi-quadric function $\Phi(r) = (r^2 + c^2)^{-1/2}$

In actual application, we have trainings on mathematical model of the network that the formula (1) described using few samples and then determine the number of parameters that being used, connection weights and proper center, the details has been presented by Guo-Hui *et al.* (2008).

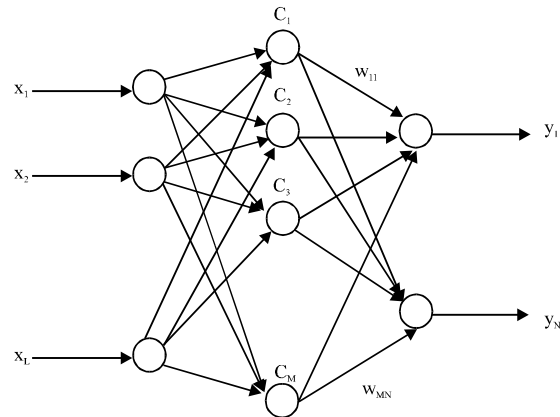


Fig. 1: RBF neural network structure

FRAME OF THE METHOD

Taking the idea of partition algorithm for reference, we first decompose the disparity image into some regions by the image processing technology, then the points clouds are optimized in every regions by the RBF neural network algorithm and then the reconstruction procedure is adequately conducted using the optimized cloud points, finally the all regions are spliced together and the integrated results of reconstruction are obtained.

Region segmentation of disparity image: The gray of every point in disparity image reflects directly the depth from the real scenery to image plane, so we also call it depth image. The depth information on real scenery is continuous and dense the changes in the gradient of the boundaries is not significant in different regions in disparity image. So, it is difficult to achieve the ideal effect in region segmentation using the conventional edge detection algorithm. Therefore, a new method integrating the selective-threshold segmentation with the conventional edge algorithm is proposed to obtain the decomposed regions. The steps are as follows:

- **Step 1:** First, estimate roughly the maximum A and minimum B of the gray values in the regions to be detected and set the original threshold range as [A,B]
- **Step 2:** Then, search for the regions those are in accordance with the threshold range in the whole image, and make fine adjustment to A and B gradually until the best results are achieved. The judgement of best results is that the region edge is full close
- **Step 3:** The candidate regions can be more than one in step 2. The regions to be detected further are to be kept and others are to be removed
- **Step 4:** Carry out conventional edge detection on the regions that remains in step 3, obtain the edge information and record it timely

- **Step 5:** Repeat the procedure of step 1, 2 and 3 until the complete edge information in all regions is obtained
- **Step 6:** Fuse all the resultant edge information together and the final segmentation results are obtained

We apply the Robert operator to extract the edges in step 4. Two 4-pixel templates are needed in the Robert operator, as shown in Fig. 2.

Robert operator has the characters such as high accuracy in edge location and being sensitive to noise. Considering that the subsequent experiment is conducted on the noiseless image, we choose this operator. When taking the noise into account, we can choose the Prewitt operator or Sobel operator instead.

RBF neural network used in this study: After the procedure, several independent regions, which are the subsets of the original image, are obtained. Meanwhile, the number of the points and the complexity of the topological structure are reduced, respectively. Iteration based on RBF neural network is applied to optimize the results in these regions and the optimal point set ready for

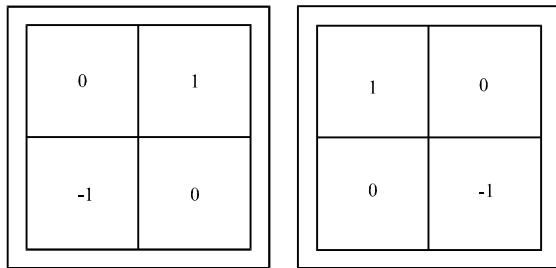


Fig. 2: Two templates of Robert operator

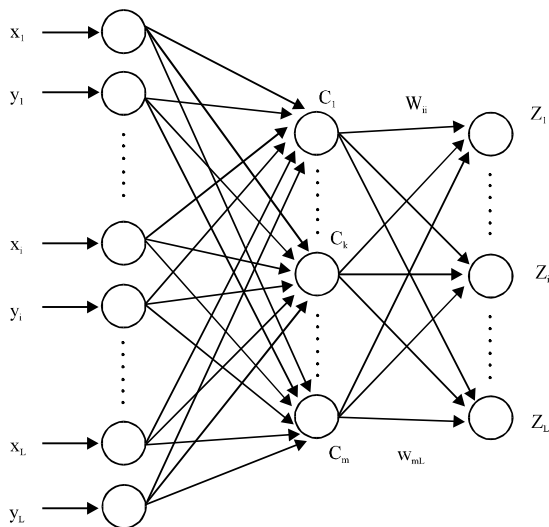


Fig. 3: RBF network used for optimizing cloud points

reconstruction is generated, which will give rise to better reconstruction effect.

When the RBF network is used to optimize the cloud points of disparity image, the network structure turns into the form as Fig. 3.

The network outputs can be calculated by the following formula:

$$Z_i = w_{i0} + \sum_{k=1}^{n_c} w_{ik} \Phi(\|x_i - c_k\|) + \sum_{k=1}^{n_c} w_{ik} \Phi(\|y_i - c_k\|) \quad (2)$$

where, x_i ($1, 2, \dots, L$) is the measured value of the i th pixel input x , y_i ($i = 1, 2, \dots, L$), is the measured value of y . c_j ($j = 1, 2, \dots, n_c$) is the center of j th radial basis function, n_c is the number of hidden layer, w_{ik} is the connection weight between hidden neuron k and output neuron i ($i = 1, 2, \dots, L$), w_{i0} is the i th threshold of target neuron, Z_i is the output of the i th target neuron Z_i to the input of the i th pixel value (x_i, y_i) , i.e., the desired output of the depth value.

In actual application, we can determine the numbers of the input layer and output layer according to the numbers of the pixels in different regions and initially determine the number of hidden layer using the empirical formula and then train the mathematical model described in formula (2) with few samples, therefore the number of RBF used in hidden layer' connection weights and the proper center are determined. The implemented method is as follows:

- **Step 1:** Pick up few samples in certain region as inputs sample and train the RBF neural network. Here we take x and y coordinate as the inputs, z coordinate as the desired output. The method of picking up samples is evenly selecting via some resolution
- **Step 2:** Determine the connection weights and the center quickly with the operation in step (1). After the definite RBF module is obtained, feed all the input data and compute the results of the approximating procedure
- **Step 3:** Carry out triangle mesh generation with the results in step (2) and get the resultant reconstruction in the regions
- **Step 4:** When RBF reconstruction is performed in the regions with a complex structure, carry out the training and approximation with the data chosen according to the original resolution of the image, here the input data set could be $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$. When RBF reconstruction is performed in the regions with a smooth structure, carry out the training and approximation with the data chosen according to the lower resolution of the image, here the input data set could be $\{(x_1, y_1), (x_3, y_3), \dots, (x_n, y_n)\}$

There is a distinct difference in the decomposed regions in the procedure above: the structures are complicated in some regions and the represented information should be large enough as to restore more true outlines; some are smooth and simple in some regions, only few represented information can restore the image with ideal results. Therefore, after the trainings on RBF network in Fig. 3, the connection weight and center in the networks of different regions are different; it represents a sequence of isomorphic RBF network.

Triangulation in the regions: Now the original disparity image has been segmented into several small regions with known boundaries and optimized cloud points. To make use of these characters, we can conduct the triangulation in the following steps:

- **Step 1:** Perform the procedure from the boundaries. Connect n points with straight lines orderly firstly and obtain $n-1$ initial sides
- **Step 2:** Take every initial boundary as a side of the triangles to be triangulated and search for the proper vertex towards the regions inward by the rules of nearest-distance, keeping the foot point of the third vertex for the triangle in line with the coordinate side and no side used for twice, as is shown in Fig. 4

Then take the sides such as BC, DE, IJ, JK as the initial sides and start the triangulation towards the regions inward. Here, for the point F is closer to the borderline DE than the point G, we firstly choose the point F to carry out the triangulation, it is the nearest distance rule; the foot point of the point H which is relative to the line IJ is

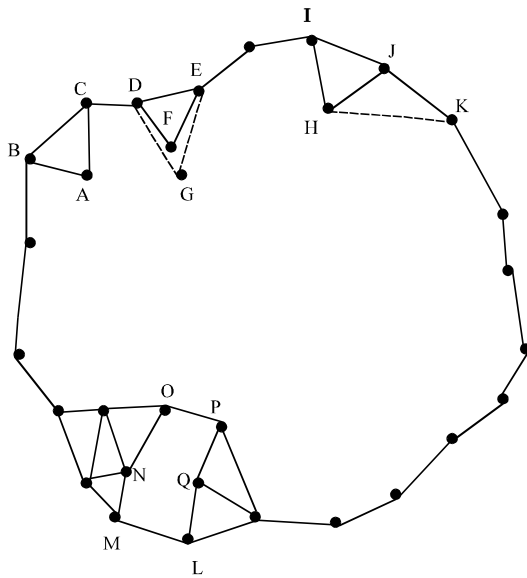


Fig. 4: Performance of triangulation

situated in the line IJ, whereas $\angle HJK$ is obtuse angle which will be abandoned for its unsatisfactory result of the triangulation, it is the keeping the foot point of the third vertex for the triangle in the line with the coordinate side rule; every resultant side can be the common side for no more than one adjacent triangle, it is the no side used for twice rule.

With these constraints described above, the triangulation is performed from the boundaries to the center until it has been done adequately. Here the results of triangulation include some regions that haven't triangulated adequately, as the curving polygon LMNOPQ is shown in Fig. 4.

- **Step 3:** Conduct the further triangulation on the curving polygon LMNOPQ. The main task is to compensate the non-triangulated regions in steps 1 and 2, so it can be relatively easy to be performed

Region splicing: From the design of isomorphic RBF neural network we can learn that the density of cloud points in different regions is diverse. The numbers of cloud points for the regions with simple structure are small while those for the regions with complicated structure are significantly larger. Thus, the numbers covered by the same boundaries in different regions are not uniform.

When the splicing is conducted, we take the grids in which the boundaries are include, as the reserve data, as is shown in Fig. 5. The structural characters are simple in the left regions and complicated in the right regions. Here we remove the boundaries in two adjacent regions, as the dotted lines shown in Fig. 5. Then redo the triangulation to splice different regions.

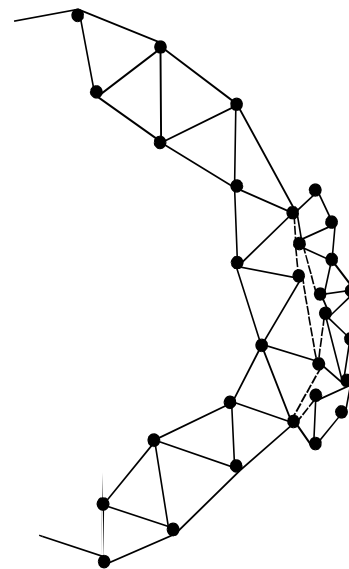


Fig. 5: Region splicing

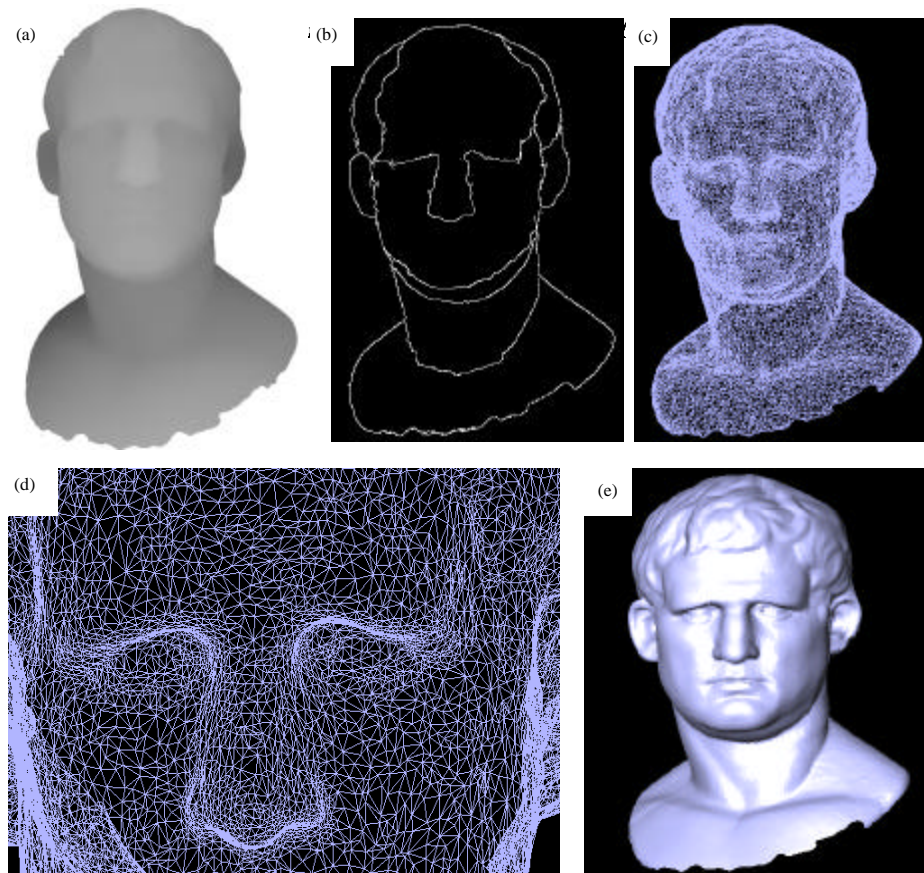


Fig. 6: Experimental results; (a) Disparity image for a head, (b) Results of region segmentation, (c) Results after triangulation (d) Magnified results of the local triangulation image and (e) Results of reconstruction

After the previous processing, now we add the light effect into the present result and the final results of the reconstruction are generated.

EXPERIMENTAL RESULTS

To testify the effectiveness of the proposed method, we conduct the experiment on the noiseless disparity image of a head. The depth information for the disparity image is continuous that it is difficult to segment it. The structural characters in different parts of the face are not uniform. The structures of the eyes, the nose and mouth are complicated; those of the facial, the forehead and neck are smooth. These characters are very targeted to testify the proposed algorithm in this study and the implemented experimental results are shown in Fig. 6.

From the Fig. 6, we can get the following information:

- The depth information in the disparity image of the head is dense and the characters in different organs are clear

- Carry out the region segmentation on Fig. 6a with the method described in section segmentation of disparity image, the results are shown in Fig. 6b. From the Fig. 6b, we can conclude that the proposed method in the paper integrated a selective-threshold decomposition with conventional edge detection, perform better segmentation results for disparity image
- Figure 6c and d is the triangulation results and the magnified results of the local triangulation image. In the processing based on isomorphic RBF network, we conduct the trainings and approximation with original resolution since the character of the eyes is most complicated; whereas the character of the nose is more complicated, we conduct the same procedure with lower resolution and the characters of the facial and forehead are smooth, we conduct the procedure with more lower resolution. Consequently, it forms the overlapping effect of the thin and dense regions in the triangulation results

- The results of reconstruction using the proposed method is shown in Fig. 6e and the surface of the reconstruction is smooth and continuous, demonstrating the excellent capacity of the non-linear approximation on the RBF network

DISCUSSION

According to previous research work, most reconstruction experiments had been done by using point cloud in 3-D space. Our research was carried out for point cloud in 2-D disparity image. So, comparison with other research had not been done. Obviously, Our research has 3 differences with other researchers:

- Our research object is not point cloud in 3-D space but point cloud in 2-D disparity image
- Our reconstruction work is in virtue of topological structures of disparity image
- Our reconstruction research combines with image processing technology

CONCLUSIONS

In this study, a reconstruction method for disparity image based on region segmentation and isomorphic RBF neural network is proposed. Several independent regions are obtained after region segmentation and the number of the points and the complexity of the topological structure in every region are reduced, respectively, which helps the reconstruction algorithm conducted smoothly. As for the reconstruction procedure in different region, we apply the isomorphic RBF neural network to optimize the cloud points and adapt to the region triangulation by trainings with corresponding resolution. Experimental results show that the proposed method can attain reconstruction results of high quality effectively.

ACKNOWLEDGMENTS

This research was supported by The Scientific and Technological Project of Education Department of Heilongjiang Province with grant No. (11541046, 2009-

2011) conducted by P.R. China, The Postdoctoral Sustentation Fund of Heilongjiang Province with postdoctoral No. (69449, 2009-2011) conducted by P.R. China, The Harbin Special Fund of Technological Innovation Talent with grant No. (2010RFQXG039, 2010-2012) conducted by P.R. China and Youth Science Research Fund Project of Harbin University of Science and Technology with grant No. (2009YF011, 2010-2011) conducted by P.R. China. National Natural Science Funds for young scholar with grant No. (50905049, 2010-2012) conducted by P.R. China and National Key Technologies R and D Program of Hei Longjiang with grant No. (GC09A524, 2009-2011) conducted by P.R. China.

REFERENCES

- Alfredo, L., L. Francesco and P. Marcello, 2010. Real-time 3D features reconstruction through monocular vision. *Int. J. Interact. Des. Manuf.*, 4: 103-112.
- Atkinson, G.A., A.R. Farooq, M.L. Smith and L.N. Smith, 2009. Facial reconstruction and alignment using photometric stereo and surface fitting. *Pattern Recognit. Image Anal.*, 5524: 88-95.
- Brovko, A.V., E.K. Murphy and V.V. Yakovlev, 2009. Waveguide microwave imaging: Neural network reconstruction of functional 2-D permittivity profiles. *IEEE Trans. Microwave Theory Techniques*, 57: 406-414.
- Grant, S. and D. Frank, 2010. Probabilistic temporal inference on reconstructed 3D scenes. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, June 13-18, San Francisco, CA, pp: 1410-1417.
- Guo-Hui, H., Z. Bin and G. Jun-Ying, 2008. 3-D face reconstruction using RBF and B spline method. *Proc. Int. Congress Image Signal Proc.*, 1: 239-243.
- Raissi, D.V., L. Fabrice and B. Benoit, 2008. Piece-wise linear DFT interpolation for IIR systems: Performance and error bound computation. *Proceedings of the 42nd Asilomar Conference on Signals, Systems and Computers*, Oct. 26-29, Pacific Grove, CA, pp: 588-592.