技術解說

3차원 화상 신호 처리 기술과 그의 응용

− 3−Dimensional Image Signal Processing Techniques & Their Application −

조동욱*, 김지영*, 유홍균**
(*서원대학교 전자계산학과,
**충북대학교 전자공학과)

Ⅰ. Introduction

The goal of 3-D image signal processing is to recognize or understand 3-D scenes from vsual input. The visual input may be a monocular image, a range image or time sequence images. A major difficulty with 3-D image signal processing is that a 3-D scene may include many objects which look differently depending on the viewer direction, illumination conditions and geometrical relations among objects. We human beings can usually understand such scenes easily by using rich sensor information acquired from an environment and knowledge about the world(or the world model). In order to realize a comparable 3-D image signal processing system, we have to solve the following problems :

- Acquiring range information of a 3-D scene
- Making a description of a scene from input images which may consist of color images and/or range images
- Interpreting the description using the world model.

On the other hand, the typical tasks of 3-D image signal processing include (a) the navigation of autonomous vehicles on the land, in the sky or under the sea (b) the assembly or inspection of manufactured parts and (c) the analysis of microscopic images and medical X-rays. In a number of applications, the goal of the 3-D image signal processing system is to identify and locate a specified object in the scene. But, the situation of the applications of 3-D image signal processing system has been quite different. In industry, which is the largest application area, the major difficulties mentioned above have been avoided by controlling the viewer direction and illumination conditions, or even by constraining the geometrical relation of objects. For example, 3-D position of an object is easily obtained from a monocular image by imposing a constraint of the object position. However, the 3-D position can not always be constrained in such a way. A typical example is flexible objects such as pipes or wires. For assembly of such objects, simple binocular vision or active range finding systems have been employed. Another important application of current 3-D image signal processing is geometrical inspection of 3-D objects. Since no geometrical constraints can be used for the inspection, range data acquisition is indispensable. The process of range data acquisition, 3-D scene representation & matching method along with the application areas of 3-D image signal processing are described in the next chapters.

Ⅱ. 3-D Image Signal Processing Techniques

An important aspect of any 3-D image signal processing system is its data acquisition module. The task is performed in one of two approaches : passive or active. In the passive approach, 3-D information is inferred from the scene using existing energy in the environment, such as reflected light. In the active approach, the 3-D information is derived by projecting external energy waves, such as sonar waves and laser light. In this chapter, the active methods are reviewed and the passive methods are dealt in the next chapter.

1. Active-Range Sensing

Active-range sensing can be divided into two main classes. In the first calss, the principles of triangulation are used. Each point in a scene is highlighted, using a sheet of light, and observed by the sensor.

Then, using the known geometry of the imaging system, the distance of each highlighted point to the sensor is caluculated (see Figure 1). Typical methods are given in the references [1]~[4].

One of the disadvantages of a first class(triangulation-based methods) is the shadow effect, where a region of the scene is not visible to either the laser or the sensor. On the other hand, in the second class(time-of-flight-range finders), a laser beam is emitted and received along the same path, eliminating the shadow problem. However, since these systems depend on the returned laser light to measure the distance, high-energy laser sources, possibly harmful to the human eye, are required. Also, most time-of-flight range finders require complex electronics, raising the

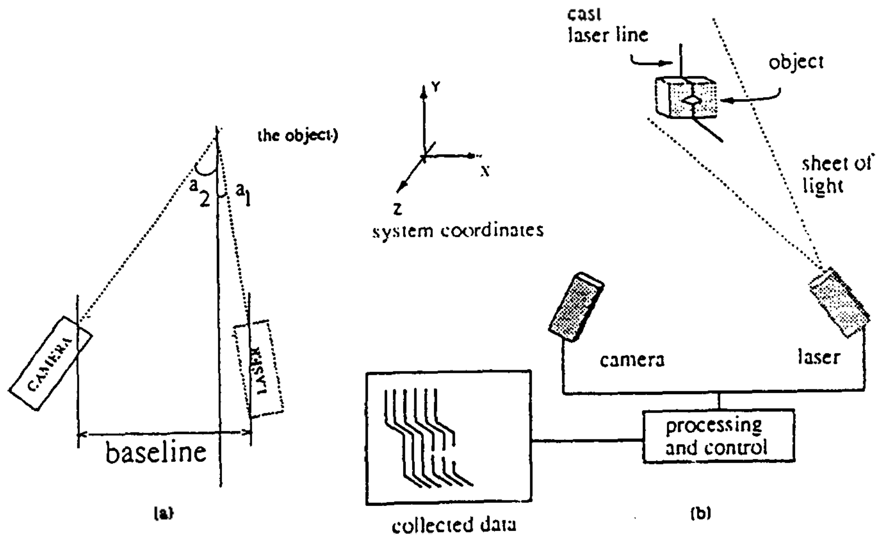

Fig 1. (a) The geometry of a typical laser triangulation system
(b) View of the setup

cost of such sensors. Two classes of laser sources are used in time-of flight scanners : pulsed [5]~[6] and continuous-beam lasers [7]~[8].

2. Segmentation of Range Images

Once the necessary measurements(e.g range images) from the scene have been made, the data must be represented using symbolic description to enable the system to carry out the specified high-level processes, such as matching. The key to performing such tasks is the means by which shapes are described by the system. In a large number of cases, local-surface properties such as surface curvature and surface normal, are used to describe the shapes. Briefly, surface curvature is the rate at which the surface derivatives has from its tangent plane. Curvature is an important measure since it is invariant to viewing directions and does not change with occlusion. However prior to the calculation of most surface properties, such as curvature and surface normal, the input data such as range images must be smoothed. The reason is twofold. First, the measurements made by the sensors are often inaccurate due to detector noise, quantization, calibration errors and other sources of error, such as "speckle". Second, the calculation of curvature involves second-order partial derivatives, magnifying the effects of any noise present. To solve the problem, Brady et al. [9] and Yang [10] have used Gaussian smoothing [11]. Once noise reduction processes (i.e smoothing) are completed, many of surface properties such as curvature and surface normal, may be approximated using one of two general approaches.

In the first method, a mathematical function, such as a spline, is fitted to the data, and then the necessary partial derivatives are solved at each desired point using the resulting function (see, for example, Vemuri et al [12]). The disadvantage of this approach is that the image must be partitioned, since no single function can approximate all area of an image. This in turn requires detection of depth discontinuities in the depth map.

The second approach to finding the surface properties uses a set of local window operators convolved with the image data points, to calculate the partial derviative at eath point. The partial derivatives are then used to approximate the desired surface properties. In this approach, the size of the operator plays an important role ; small windows might not yield good approximations to curvature because the area may be too planar. On the other hand, large window operators may overlap areas which include surface discontinuities, causing erroneous measures close to boundary points. Nevertheless, the second and more computationally efficient approach has been more popular. This is partly because most systems in the past have used only the sign of the curvature values rather than the actual values. Further detailes are found in references [13]~[18].

3. Representation

Representations are used to describe shapes. Unlike quantifiable entities such as motion and intensity, the description of complex 3-D shapes requires hundreds of parameters and is complicated [19]. Such representation is not computationally feasible today if it is intended to be used in matching two objects. However a representation may be feasible for matching if it is (a) unambiguous(no two different objects have the same representation), (b) unique(there is a single description for each object using the representation scheme), (c) not sensitive(with respect to missing data points, such as in the cases of occlusion), and (d) convenient to use, in the matching stage, and to store. This section divides representations used for 3-D object descriptions, into surface-based, discontinuity-based, and volumetric-based schemes, and reviews each class briefly.

(1) Surface-based representation

In the class of representations, surface properties, such as surface normal and Gaussian curvature, are approximated and used in the description of the collected data points. Among the surface-based representation methods, Gaussian sphere method maps the orientation of the surface normal at each point of the surface to the unit sphere [20]~[21].

The moment-based method defines the properties of shape by the moments of functions after the shape is described by mass function [22]. In the discontinuity-based representation, rather than store information about the surfaces, the information about the points where the characteristics of an area on the surface change are preserved, and the curves embedded in 3-D space are detected and represented by mathematical means such as parametric polynomial curves or rational B-splines.

Other schemes include the space curve method which descrives bounding contours [23] and the surface primal sketch method. The surface primal sketch method extends curvature primal sketch [24] by detecting and classifying several types of discontinuities, such as steps, roofs, smooth joins, shoulders, etc. into model [25]. The aspect graph method represents each distince 2-D, viewpoints of 3-D objects as a node, transformations and visual events as arcs [26]. In the discontinuity labeling method [27]~[28] similar to polyhedral edge junction graph, each node represents polyhedral vertex, curved junction, self-occlusion, etc., and link eack represents convexity, concavity, occlusion, limb(defined as a depth discontinuity by a smooth surace curving away from the viewer), etc.

(2) Volumetric Representations

This class of representations describe volumes rather than surface or discontinuities of surfaces. While most such representation schemes are efficient in describing shapes, their disadvantage is that first, in most cases, the objects have to be symmetrical and simple in shape. Second, these representations may not be used directly in the matching stage. The class of representations has been used most often for CAD modeling rather than for recognition tasks. Volumetric representation method includes super quadrics [29], octrees [30], CSG(Constructive Solid Geometry) [31], sweep representation [32], etc.

4. Matching

Once the appropriate descriptions are derived from the data and the appropriate model, the 3-D image image signal processing system is able to match the two descriptions completing the task of object recognition. This is performed in two steps. In the first step, a correspondance is established between the two sets of descriptions. The matching strategy must be able to achieve this step using a partial description of the object and its full model description to account for possible missing data points due to the possibility of partial occlusion, a single viewpoint, field of view, etc. Mostly used methods for this processing are tree search method [33]~[35] and attributed graphs method [36]~[38]. In the second step, using the established correspondances, a geometrical transformation(usually a ratation matrix and a translation vector) is derived such that the model may be transformed to the orientation of the object in the scene. For this, Ikeuchi & Hong [39] established a free coordinate system, and Faugeras & Hebert [40] used quaternions to represent rotation matrix. Also the generalized Hough transform [19] is a very useful way in orientation determination.

Ⅲ. Applications

In the 3-D shape inspection, since geometrical constraints cannot be applied, range data become indispensable. Active methods(i.e. range finder) are adopted in the indoor applications and for the outdoor use passive methods(i.e. stereo vision) are applied.

1. 3-D Shape Inspection

Most of practical 3-D vision sensors are based on active triangulation ; they project a light beam and detect the reflected light with an image sensor. Although some of them were developed for recognition of 3-D objects, a few has been actually used for recognition. Toyota Central R&D Lab., for example, developed a 3-D vision sensor which can obtain range data of 150 × 241 points in 2.5 sce [41].

It consists of a laser for projecting a sheet of light, a TV camera for obtaining the reflected light image, and a processor for determining the position of the light(see Figure 2).

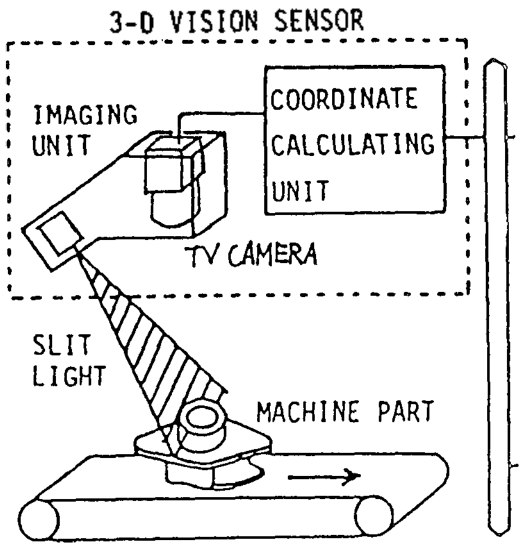

Fig 2. A 3-D vision sensor for object recognition

The initial goal was classfication of machine parts which can not be discriminated from the outer contour alone. The classifier uses, in addition to the outer contour, the shape of a few sections at specified heights which is derived from the range data. Currently, it is used for inspection of 3-D positions of assembled machines, the 3-D shape of machine parts such as body panels [42], cylinders or IC chip leads [43].

2. High Speed Streo Vision for Autonomous Vehicle

Although the active triangulation range finder described in the previous chapter is reliable in controlled environments, it is not suitable for use in outdoor environ-ment because the illumination is very strong in daytime and the light relfected on distant objects is too weak. Passive stereo vision has long been studied to be used in such a case. Major difficulties in stereo vision are the following :

- There is no established method for finding correct correspondance between a stereo pair of images
- Finding correspondance requires much computation

For simple environments, however, those problems can be avoided. An example is a road scene where only obstacles are big objects such as cars. Tsugawa et al [44] made a wired circuit stereo vision for an autionomous vehicle which ran on a test course with guard rails.

Recently many companies have interest in stereo vision for obstacle detection in outdoor environments as well as 3-D geometrical inspections in factories. An example of the former application is the assistance of a driver by finding other cars on the way. For practical use, simpler stereo vision is developed to achieve a quick response. Based on the principle similar to [44], a stereo vision system is realized by a variable structured image processor for "Personal Vehicle System" developed by Fujitsu and Nissan. This system at first finds white lines on a road and them searches the candidate area determined from the white lines for obstacle [45]. A pair of cameras are tiled 90 degrees and mounted between two front lamps at different heights. The horizontal scanning of both cameras, therefore, corresponds to a vertical scanning in the actual 3-D space. The video signals are differentiated and edges are detected by thresholding. Note that horizontal edges are easily detected by the vertical scanning, while vertical ones or steeply tilted ones are not.

Assuming that obstacles are brighter than dark asphalt road, and that they have approximately horizontal edges, the principle of the stereo is to find a positive and negative edge pairs in a stereo pair of images during each scanning period, and obtain a corresponding pair of the edge pairs as shown in Fig. 3.

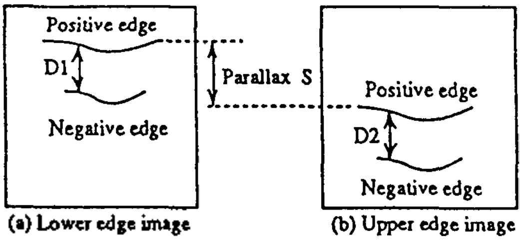

Fig 3. Principle of a simple stereo matching

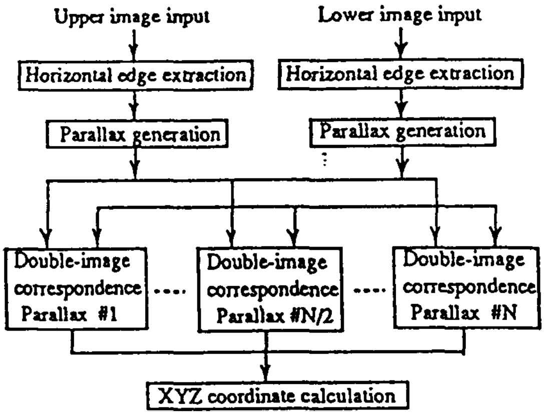

Fig 4. Stereo matching by a variable structured image processor

Actually the variable structured image processor first finds corresponding edges in a pair of images, and then another computer obtains positive and negative pairs. In order to obtain corresponding edge, the processor generates multiple pairs of image with different parallax as shown in Fig. 4, and finds matching pairs of edges for each pair of images.

Among the obtained matched edges, positive and negative edge paris are selected These edge pairs are classified, according to the distance computed from the corresponding parallax, into obstacle candidates or just edges on the road. Only when enough number of connected obstacles are found in consecutive scannings these candidates are determined to constitiute an obstacle. This stereo vision system worked for finding cars in real time on a clean test course. The system, however, is not reliable enough for actual use. Two of the important limitations are the followings :

- Only distinct edges are used for finding the stereo pairs of correspondance
- That only a pair of positive and negative edges is detected as an obstacle candidate

Ⅳ. Conclusion

A 3-D image signal processing system consists of a range finding process to acquire range data, range image segmentation, representation process & matching process. The processed range data find many application areas in 3-D shape inspection. Further, 3-D image signal processing techniques are used in ALV, in reconstructing 3-D scene structures [46] where color and range information are combined in the process, in 3-D information extraction [47] where monocular colour images undergo scene interproetation process. However, object recognition is not widely used in industry. The speed of current systems, from data collection to identification, is slow for most industry applications.

Segmentation and low-level image-processing tasks in general are the major bottleneck in the process. However, some of the existing methods could be perfomed using parallel processing techniques to reduce the required time. So far, even though many methods are proposed and used, each method has its own drawback and no method is proven to be superior in image recognition for all circumstances. It is because there is not a well-defined function to model the human vision system. It is believed the reason comes from the fact that all the traditional approaches are based entirely on the crisp methods. This lead the research efforts to new directions, such as in fuzzy, neuro and neuro-fuzzy system, for better image processing. In the future, many new techniques in this closer-to-human direction are expected to be suggested and parallel processing will play important part to increase the image processing speed.

참 고 문 헌

1. Besl, P. J., "Active optical range imaging," In Advances in Machin Visio, J.L.C. Sanz Ed. Spriger-Verlag, New York, 1989.
2. Nitzan, D., "Three-dimensional vision structure for robot applications," IEEE Trans. PAMI, 291~309, 1988.
3. Freeman, H., "Machine Vision Algorithms, Architectures and Systems," Academic Press, San-Diego, Calif., 1988.
4. Kanade, T., "Three-Dimensional Machin Vision," Kluwer Academic Pubishers, 1987.
5. Lewis, R. A. & Johnston, A. R., "A scanning laser range finder for a robotic vehicle," Proceedings of the 5th IJCAI, pp. 762~768, 1977.
6. Jarvis, R. A., "A Perspective on range finding techniques for computer vision," IEEE Trans. PAMI, pp. 122~139, 1983.
7. Jarvis, R. A., "A laser time-of-flight range scanner robot vision," IEEE Trans. PAMI, pp. 505~512, 1983.
8. Fu, K. S, Gonzalez, R. C., & Lee, C. S. G., "Robotics : Control, Sensing, Vision and Intelligence," McGraw-Hill, 1987.
9. Brady, M., Ponce, J., Yullie, A. & Asada, H. "Describing surfaces," CVGIP, pp. 1~28, 1985.
10. Vang, H. S., "Range image analysis via quadtree & Pyamid structure based on surface curvature," In Intelligent Robots & Computer Vision : Seventh in Series, D. P. Casasent, ED. Proceedings of SPIE, pp. 597~608, 1988.
11. Marr, D. & Hildreth, E., "Theory of edge detection," In Proceedings of the Royal Society of London, pp. 195~240, 1980.
12. Vemuri, B. C et al, "Curvature-based representation of objects from range data," Journal of Image Vision Computer, pp. 107~114, 1986.
13. Hoffman, R. & Jain, R. C., "Segmentation and classification of range images," IEEE Trans. PAMI, pp. 608~620, 1987.
14. Sabata, B. et al, "Segmentation of 3-D range images using pyramidal data structures," Proceedings of the International Conference in Computer Vision, pp. 662~665, 1990.
15. Boulanger, P. et al, "Segmenation of planar and quadratic surfaces," In Intelligent Robots and Computer Vision : 6th in Series. Proceddings of SPIE848, pp.395~403, 1987.
16. Han, J. et al, "Range image segmentation and surface parameter extraction for 3-D object recogintion of industrial parts," Proceddings of the 1987 IEEE Conference on Robots & Automation, pp. 380~386, 1987.
17. Flynn, P. J. et al, "Surface classification : Hypothesis testing and planar estimation," Proceedings of CVPR, pp. 261~267, 1998.
18. Taylor, R. W. et al, "Fast segmentation of range imagery into planar regions," CVGIP, pp. 42~60, 1989.
19. Ballard, D. H. & Brown, C. M., "Computer Vision," Proetice-Hall, 1982.
20. Hilbert, D. & Cohn-Vossen, S., "Geometry and the Imageination," Chelsea, New York, 1952.
21. Horm, B. K. P., "Robot Vision," MIT Press, 1986.
22. Taubin et al, "Representing and Comparing shapes using shape polynomials," Proceddings of CVPR, pp. 510~516, 1989.
23. Mokhtarian, F., "Multi-scale description of space curves and three-dimensional objects," Proceedings of CVPR, pp. 298~303, 1988.
24. Asada, H. & Brady, M., "Curvature primal sketch," IEEE Trans. PAMI, pp 2~4, 1986.
25. Ponce, J. et al, "Toward a surface primal sketch," Three-Dimentional Machine Vision, T. Kanade,

Ed. Kluwer Scsdemic Publishers, pp. 95~24J, 1987.

26. Koenderik, J. J. et al, "The internal representation of solid shape with respect to vision," Bio. Cybernetics 32, pp. 211~216, 1979.

27. Godin, G. D. & Levine, H. D., "Structured edge map of curved objects in a range image," Proceddings of CVPR, pp. 276~281, 1989.

28. Chen, S. & Stockman, G., "Object wings-2 1/2-D Primitives for 3-D Recognition," Proceddings of CVPR, pp. 535~540, 1989.

29. Entland, A. P., "Recongition by parts," Procedings of the 1ts International Conference in Computer Vision, pp. 612~620, 1987.

30. Mazumder, P., "A new strategy for octree representation of three-dimensional objects," Proceedings of CVPR, pp. 270~275, 1988.

31. Foley, J. D. et al, "Fundamentals of Interactive Computer Graphics," Addison-Wesley, 1984.

32. Binford, T. O., "Visual perception by computer," IEEE Conference on System & Control, 1971.

33. Grimson, W. E., L., "On the recognition of curved objects," IEEE Trans. PAMI, pp. 632~642, 1989.

34. Grisen, W. E. L. et al, "Localizing overlapping parts by searching the interpretation tree," IEEE Trans. PAMI, pp. 469~482, 1987.

35. Grimson, W. E. L., "On the recognition of parameterized objects," 4th International Symposium on Robotics Reseach, 1987.

36. Fan, T. J., "Describing and Recognizing 3-D Objects Using Surface Properties," Springer-Verlag, New York, 1990.

37. Fan, T. J. et al, "Recognizing 3-D objects using surface descriptions," IEEE Trans. PAMI, pp. 1140~1157, 1989.

38. Fan, T. J. et al, "Segmented description of 3-D surface," IEEE int. J. Robot Automat. pp. 527~538, 1987.

39. Ikeuchi, K. & Hong, K. S., "Determining linear shape change : Toward automatic generation of object recognition pograms," Proceedings of CVPR, pp. 450~457, 1989.

40. Faugeras, O. D. & Hebert, M., "The representation, recognition, and positioning of 3-D shapes from range data," Techniques for 3-D Machine Perception, A. Rosenfeld, Ed. North-Holland, Netherlands, pp. 13~52, 1987.

41. Yamamoto, S. et al, "Recognition System for Machine Parts with a 3-D Vision Sensor," Proc. SPIE, pp. 698~702, 1987.

42. Higuchi, K. et al, "Inspection of Autmotive Body Panels Dimension," JIEICE TR PRU, vol. 90, No. 51, pp. 79~85, 1991.

43. Tsukada, T. et al, "Idem," vol. 90, No. 412, pp. 33~38, 1990.

44. Tsugawa, S. & Yatabe, T., Hirose, T. & Matsumoto, S., "An Automobile with Artificial Intelligence," Proc. 6th IJCAI, pp. 893~895, 1979.

45. Ohzora, M. et al, "Wideo-Rate Image Proceeding System for an Autonomous Personal Vehicle System," Proc. IAPR Workshop on Machine Vision Applications, pp. 389~402, 1990.

46. Okamoto, A., Shirai, Y. & Asada, M., "Reconstruction of 3-D scene Structure by Integrating Color and Range Information," Proc. Korea-Japan Computer Vision Conference, pp. 77~80, 1990.

47. Hirata, S. Shirai, Y. & Asada, M., "Scene Interpretation using 3-D Information Extracted from Monocular Color Images," Proc. Korea-Japan Computer Vision Conference, pp. 81~87, 1991.

▲Dong-Uk Cho(Regular Member)

He received the B.S, M.S and Ph.D in Electronic Engineering all from the Hanyang University in 1983, 1985 and 1989 respectively. He is an assistant professor of computer science and engineering at Seowon University, Chongju since 1991.

Before joining Seowon University faculty, he was an assistant professor of telecommunication engineering at Dongyang Technical College for 2 years.

His research interests include computer vision, fuzzy set theory and neural network. He is a member of the IEEE PAMI and KITE.

▲Heung-Gyoon Ryu(Regular Member)

12권 2E호 참조

▲Ji-Yeong Kim(Regular Member)

He is an associate professor of computer science and engineering at Seowon University, chongju since 1989 where he has been also the Director of the University Computer Center. His research interests include system performance evaluation, computer vision and tele-measuring and controlling system.

Before joining Seowon University faculty, he was the Chief Research Scientist at the Central Research Center at Oriental Precision Company Ltd. for 2 1/4 years. Before then he was an assistant professor of computer science and engineering at Auburn University, Auburn, Alabama for year.

He received M.B.A in MIS and M.S. [illegible] in computer science all from the State [illegible] of New York at Binghamton NY. [illegible] and 1984 respectively. He is a member [illegible] IEEE Computer and Communication [illegible] and the ACM.