2003－ＡＶＭ－43　(20)
2003／12／19

マルチユーザのためのカメラセンサ・ネットワークの最適化

メヒルダドパナヒプルテヘラニ† 藤井俊彰† 谷本正幸†

†名古屋大学大学院 工学研究科電子情報学専攻谷本研究室
〒464-8603 名古屋市千種区不老町
E-mail: †mehrdad@tanimoto.nuee.Nagoya-u.ac.jp, †{fujii, tanimoto}@nuee.nagoya-u.co.jp

あらまし 本稿は,任意視点生成のための,マルチユーザ用カメラセンサー・ネットワークの最適化について検討した．シングルユーザの場合の処理時間と通信量については,各ノードが任意視点生成アルゴリズムの処理を分散することで最適化することができる．そこで，我々は共有された画像サイズに基づいた２つの方法，F-DP(Fully image shared Distributed Processing)，と P-DP(partially image shared Distributed Processing)を用い， それらを CP(Centralized Processing) と比較した．また， シングルユーザを拡張して,マルチユーザを想定した場合，２つのカテゴリに分類することができる．１つは，要求される視点が，異なるノードの画像を必要とする場合で，処理はユーザごとに独立しているため，シングルユーザの場合と同様に考えることができる．もうひとつは，要求される視点が，共通のノードの画像を必要とする場合であり，処理は各ユーザに依存する．よって,後者の場合について,我々はネットワークの最適化方法を提案する．最適な方法(CP, P-DP, F-DP)は、ネットワークの通信遅延と，ノードの処理能力に基づいて検討を行った．

キーワード マルチユーザ，最適化，カメラセンサ・ネットワーク

Optimization of Camera Sensor Network for Multiuser

Mehrdad PANAHPOUR TEHRANI† Toshiaki FUJII† Masayuki TANIMOTO†

†Department of Information Electronics, Graduate School of Engineering, Nagoya University
Furo-cho, Chikusa-ku, Nagoya, 464-8603, Japan
E-mail: †mehrdad@tanimoto.nuee.Nagoya-u.ac.jp, †{fujii, tanimoto}@nuee.nagoya-u.co.jp

Abstract The optimal operation of camera sensor network for multiuser case has been considered for arbitrary view generation. For optimization of processing time and communication amount in single user case, e ach node processes part of arbitrary view generation algorithm. We use two methods based on the image size shared, which are F-DP (Fully image shared Distributed Processing), and P-DP (Partially image shared Distributed Processing). They were compared with CP (Centralized Processing) method. To extend the single user case to multiuser; it can be categorized into two cases. One is independent user where the requested viewpoints need different nodes images. It is similar to single user case. Another is dependent user, where users' requested viewpoint needs common nodes image. In this later case the network should be optimized. T he best processing method (CP or P-DP or F-DP) can be chosen based on communication delay in network and the processing ability of nodes for optimum number of nodes.

Keyword Multiuser, Optimization, Camera Sensor Network

1. Introduction

We have proposed FTV (Free Viewpoint Television) system [1] where users can freely control the viewpoint position of any dynamic real-world scene. This system can cover a limited space. To increase the coverage to a wider area, distributed sensor network can be used. In a paradigm of Smart Dust, hundreds or thousands of sensor nodes of cubic millimeter dimension can be scattered about any desired environment [2]. It is used for such tasks as surveillance, widespread environmental sampling, security, and health monitoring and it is known to make human life safer and easier. We then introduce CSN (Camera Sensor Network) as an extension to FTV system, and which has hundreds of cameras, distributed densely throughout the environment. CSN presents a significant trade-off between powers consumed by processing versus communication. Communication power costs can vastly exceed today's power-efficient processor demands. As a result, in general, developers then strive to process information locally to reduce the data transmitted. The distributed architecture was designed to help sensor network capitalize on the collective behavior of these complex systems by dynamically increasing communication load only when doing so is optimal. Except for data fusion in CSN, we want to generate arbitrary view requested from users. The processing tasks in the network are capturing and calibration, rectification, corresponding search and weighted averaging to generate a new view. The processing architecture is specified based on the network processing tasks. Capturing, calibration and weighted averaging can be performed independently at any node. Rectification requires a small amount of communication among neighbor nodes. Corresponding search needs high processing cost therefore DP (Distributed Processing) is needed for arbitrary view generation. Several arbitrary view generation methods have been proposed such as pixel based, block-based [3] and offset block based [4] methods during matching algorithm for correspondence search. Arbitrary view generation in CSN can be done using CP or DP method. In CP, all processing is done in one node. Based on the image size shared, we proposed two DP methods [5], which are F-DP and P-DP.

In General, using CSN, many users can access to processed data. In this paper, we have proposed an optimized distributed processing method in CSN for case of multiuser. We categorize multiuser case into two cases. One is IU (Independent User), where the requested viewpoints need different nodes images. Another is DU (Dependent User), where users' requested viewpoints need common nodes images. The proposed optimization process decides the optimum number of nodes for multiuser case. Furthermore, the best processing method (CP or P-DP or F-DP) can be chosen based on communication delay in network and the processing ability of nodes for optimum number of nodes.

2. Camera Sensor Network

2.1. Nodes and Users Configuration

In a CSN, cameras are installed in a dense configuration. Each camera is connected to an individual processor and each node is able to communicate with other nodes using a transmitter and a receiver. Each camera with processing and communication parts is called a camera SN (Sensor Node). All nodes, which are in a same cluster, can communicate with each other. For each cluster, a CN (central node) is assigned for user interface and network management. The central node tasks are as follow. (1) to receive the location of new view from user. (2) to determine which nodes will capture the image. (3) to decide the optimum processing method (i.e. CP or DP) (4) to initiate the distributed processing among the assigned sensor nodes and (5) to deliver the generated view to each user. Note that the communication between central node and users (i.e. user interface "T_{ui}") can be optimized using the common VoD (video-on-demand) methods [6].

Multiuser can be categorized into two cases based on the location of each user. One is independent user, where the requested viewpoints need different nodes images. It is similar to single user (i.e. SU) case. Another is dependent user, where users' requested viewpoints need common nodes images. Fig. 1 shows multiuser (i.e. MU) categorization according to the location of users. In this figure, we assumed an array of four camera sensor nodes on a line and two users (u1 and u2).

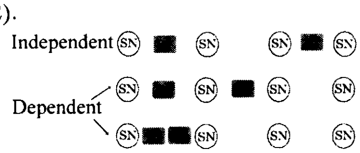


Fig. 1: Multiuser categorization based on location of users

2.2. Processing Task

The processing tasks in CSN are capturing, calibration, rectification and arbitrary view generation. Arbitrary view generation includes corresponding search and weighted averaging to generate the new view. Generating an intermediate view image "I" from the multiview images is carried out by ray-space interpolation. This is the same process as that of increasing the vertical resolution of the EPI (Epipolar Plane Image) lines constructed from multiview. Generally, interpolation is used to up-sample the EPI lines and to find the best disparity for corresponding pixels. The filter that corresponds to the found disparity is applied to up-sampled ray-space data [7] and then generates the new view. The interpolation of ray-space is shown in Fig. 2.

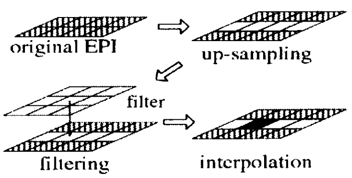


Fig. 2: Ray-space interpolation using adaptive filtering.

Generally when the multiview images are captured from different cameras, the accurate and fast interpolation between two views requires that the images are rectified [8]. The only required data for is the pair of PPM (Perspective Projection Matrices) of the two camera nodes, which can be estimated using many existing calibration methods. Interpolation can be performed directly after rectification of image pairs. Depending on the arrangement of cameras in CSN, one or two steps of interpolation are needed. Cameras can be installed in either one or two-dimensional arrays. If the cameras are installed on a one-dimensional array, for each requested viewpoint located on the line passes through optical centers of each pair of cameras, one-step interpolation is done. However, in the case of two-dimensional array, two steps interpolation is required. In two-dimensional case, the new viewpoint is assumed to be located on the three cameras plane. The camera plane passes through the three cameras' optical center. As it is shown in Fig. 3, to generate "**I**" view having C1, C2, and C3 (three cameras) images, first we need to generate interpolated image "**I1**" view with captured images from C1 and C2. Then, "**I**" view is generated by interpolation between "**I1**" and captured image from C3.

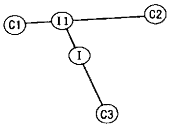


Fig. 3: Arbitrary view generation in CSN

3. Processing Methods

3.1. CP Method

In CP method as shown in Fig 6a, all the sensor nodes send the observed information to one sensor node for each user. In this method, one sensor node will be responsible for generating the arbitrary view. The generated viewpoint will be sent to the central node so that it can be delivered to the user.

3.2. DP Method

In DP method as shown in Fig. 6b, to share the corresponding search among the nodes for each user in CSN, some preprocessing and reprocessing tasks are needed before and after corresponding search, respectively. In preprocessing the to be shared image data of the SN1 and SN2 are determined and sent to "*n*" nodes (SN1 to SNn) for corresponding search. In reprocessing, SN1 collects the processed data from all other nodes to generate the image to be interpolated. The generated images are sent to the central node so that they can be delivered to the user. Next, two methods for DP have been proposed based on the shared image size.

3.1.1. F-DP Method

In the F-DP method, the searching area to find the best disparity is divided. As a result, all nodes need full image data. In preprocessing, rectified images of SN1 and SN2 are fully shared among "*n*" sensor nodes. The correspondence search for the best disparity is done in each node for "*1/n*" of maximum disparity as shown in Fig. 4. After performing the corresponding search, all the "*n*" sensor nodes will transmit the found disparity in their searching area, and its MSE (Mean Square Error) to SN1. For the reprocessing, SN1 chooses the disparity with the minimum MSE value among all received processed data, and applies weighting average to its found disparity.

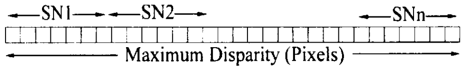


Fig. 4: Searching area (pixels) of each SN for F-DP.

3.1.2. P-DP Method

Using this method, the captured images by SN1 and SN2 are vertically divided into "*n*" parts and shared among "*n*" SNs after rectification (i.e. preprocessing) as shown in Fig. 5. The interpolation will be done in each node for "*1/n*" part of an image. The interpolated pixels value for "*1/n*" of the requested view will be transmitted to SN1. In reprocessing, SN1 generates the interpolated view by adding "*n*" parts to an image plane.

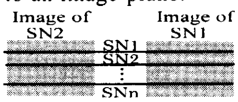


Fig. 5: Processing of P-DP method in each SN.

4. Multiuser Analysis

4.1. Assumptions

For simplification, in addition to assumptions we made in 2.2, we assume that the communication delay between all pairs of nodes does not depend on the distance between two nodes, and all users' requested view resolutions are the same.

4.2. Processing and Communication

The total processing time in case of single user "T_{su}" to generate an arbitrary view depends on number of nodes, processing ability of each node and communication delay in network. "T_{su}" is summation of total communication delay "t_d" and processing time in nodes "T_P". "T_P" is independent of network configuration. It is a function of number of nodes, interpolation time "T_i" for an image "I" and the total preprocessing and reprocessing times "T_{prp}". "T_{prp}" is independent of interpolation method. "t_d" is total network communication delay for each arbitrary view generation and depends on number of nodes and communication delay of an image "t_I".

Eq. (1) shows "T_{su}", normalized to interpolation time "T_i" where $x = \frac{t_I}{T_i}$ and $y = \frac{T_{prp}}{T_i}$.

$$\overline{T_{su}} = \overline{t_d}(n,x) + \overline{T_P}(n,y) \quad (1)$$

x and y correspond to communication and processing abilities of the network, respectively. To formulate "T_{su}" the following parameters are needed.

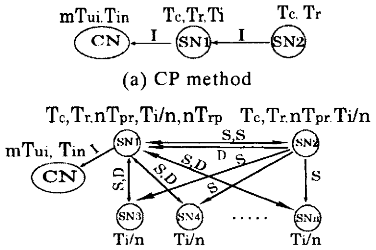
"k" is the number of interpolation (k=1,2).

"t_S" is communication delay to send the shared data "S" to "n-1" other nodes after preprocessing.

"t_D" is communication delay for processed data to send to one node from "n-1" other nodes for reprocessing.

"T_{prp}" is summation of preprocessing "T_{pr}" and reprocessing "T_{rp}" times.

"T_0" is a common processing time for all methods, which includes initiating time of SNs and defining their tasks "T_{in}", capturing time "T_c", rectification "T_r", and delay time to transmit the interpolated image to central node "t_I".



(a) CP method

(b) DP method

Fig. 6: Transmitted data and processing time in DP for $k=1$.

Fig. 6 depicts the communication amount, and processing time of each node for CP, F-DP, and P-DP, when the image to be interpolated is located on the line, passing the two cameras' optical centers in SN1, and SN2.

According to Fig. 6, "t_d" as a function of "n" $(n \geq 2)$ and "t_I", and can be shown as in Eq. (2).

$$t_d(n,t_I) = \begin{cases} kt_I & CP \\ k\left(nt_S + (n-1)t_D\right) & DP \end{cases} \quad (2)$$

"T_P" as shown in Eq. (3) is a function of "n", and processing time in nodes.

$$T_P(n,T_i) = \begin{cases} kT_i + T_0 & CP \\ k\left[\frac{T_i}{n} + nT_{prp}\right] + T_0 & DP \end{cases} \quad (3)$$

Hence, the normalized total processing time "$\overline{T_{su}}$" can be shown as Eq. (4), where $z = \frac{T_0}{T_i}$.

Total processing time for multiuser case "T_{mu}" normalized to "T_i" is shown in Eq. (5) for DU. It is summation of the total processing time for single user "T_{su}" and a linear function of x.

$$\overline{T_{su}} = \begin{cases} k(x+1) + z & CP \\ k\left[2x + \frac{1}{n}(1-x) + ny\right] + z & P\text{-}DP \\ k\left[n(2x+y) - x + \frac{1}{n}\right] + z & F\text{-}DP \end{cases} \quad (4)$$

$$\overline{T_{mu}} = \overline{T_{su}}(n,x,y) + m_u x \quad (5)$$

where "m_u" is users' coefficient and depends on number of users and users' location and can be calculated using Eq. (6) and Eq. (7). To understand the equations, in Fig. 7, we categorized DU into C-DU (Cascaded Dependent Users) and S-DU (Separated Dependent Users).

$$m_u = k \underset{j=1}{\overset{u_{SD}}{Max}} \left\{ \underset{i=1}{\overset{u_{CD}}{Max}}(a_{2i+1}) + \underset{i=1}{\overset{u_{CD}}{Max}}(b_{2i}), c_j \right\} \quad (6)$$

$$\begin{array}{lll} CP: a_i = b_i = 2m_i^{CD} - 1, & & c_j = 2(m_j^{SD} - 1) \\ DP: a_i = m_i^{CD} - 1, & b_i = a_i + 1, & c_j = m_j^{SD} - 1 \end{array} \quad (7)$$

"m_i^{CD}": Number of C-DU between ith "two views".

"m_j^{SD}": Number of S-DU between jth "two views".

"u_{CD}": $Max(i)$. "u_{SD}": $Max(j)$. Note that $m_u^{CP} > m_u^{DP}$.

The "two views" can be each pair of "C1", "C2" "C3" and "I1" of Fig. 3. "i" and "j" are the indexes refer to each pair with user.

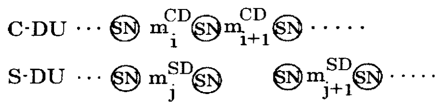


Fig. 7: Categorization of DU

4.2.1. Optimization of n

n_{op} can be calculated by solving $\frac{\delta \overline{T_{mu}}}{\delta n} = 0$ using Eq. (5). Eq. (8) shows n_{op} for P-DP, and F-DP.

$$n_{op} = \begin{cases} \left(\frac{1-x}{y}\right)^{1/2} & P\text{-}DP \\ \left(\frac{1}{2x+y}\right)^{1/2} & F\text{-}DP \end{cases} \quad (8)$$

It can be seen that the optimum number of nodes does not depend on users' configuration according to our assumptions in section 4.1. Using Fig. 8, the optimum number of nodes can be obtained. These graphs are general and can be used for any CSN.

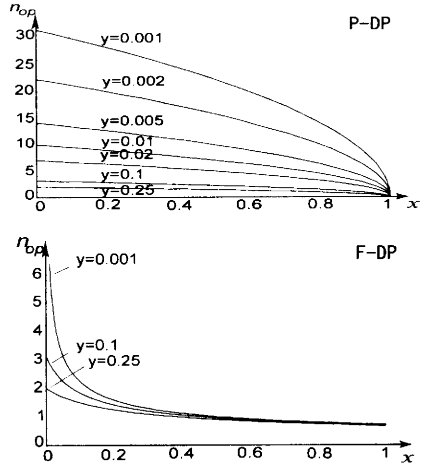


Fig. 8: Optimum number of nodes in DP vs. x and y

In the case where the optimum number of nodes for DP is less than 2, CP method should be used. Eq. (9) shows the relationship between x and y when the optimum number of nodes in DP is more than 2.

$$\begin{cases} 4y + x \leq 1 & P\text{-}DP \\ 4y + 8x \leq 1 & F\text{-}DP \end{cases} \quad (9)$$

Except number of nodes, another important factor for optimal operation of network is user's distribution. If all users, or the number of users per each "two views" are distributed homogonously among nodes the least processing time can be obtained. If "m" is total number of node, the total node "n_t" for optimal operation is "mn_{op}".

4.2.2. Comparison of CP and DP

To detrmine the best processing method, CP and DP must be compared by solving $T^{CP}=T^{DP}$ using Eq. (5). Eq. (10) and Eq. (11) show these values as a function of x and y.

$$T_{mu}^{CP} = T_{mu}^{P-DP} \rightarrow y = \left[1+\left(m_u^{CP} - m_u^{DP} - 1\right)x\right]^2 \Big/ (4-4x) \quad (10)$$

$$T_{mu}^{CP} = T_{mu}^{F-DP} \rightarrow y = \left[1+\left(m_u^{CP} - m_u^{DP} + 2\right)x\right]^2 - 2x \quad (11)$$

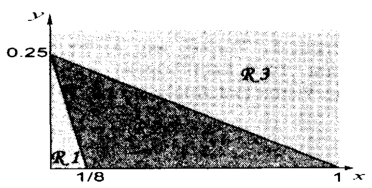


Fig. 9: Processing method decision graph

The best processing method is determined based on the region of x and y in the graph of Fig. 9. These boundaries in the graph are drawn using Eq. (9). The boundary values of Eq. (10) and Eq. (11) are not shown in Fig. 9 for simplification. The best processing method can be decided in each region using Eq. (10) and Eq. (11). A ranking of the processing methods in $\mathcal{R}1$, $\mathcal{R}2$, and $\mathcal{R}3$, is shown in Eq. (12). As it can be seen, in $\mathcal{R}1$ all processing methods can be compared, but in $\mathcal{R}2$ F-DP and in $\mathcal{R}3$, comparisons between both both P-DP and F-DP are not valid to be compared.

$$\begin{cases} \mathcal{R}1: & T_{P-DP} < T_{F-DP} < T_{CP} \\ \mathcal{R}2: & T_{P-DP} < T_{CP} \\ \mathcal{R}3: & T_{CP} \end{cases} \quad (12)$$

5. Experiment

In Fig. 10, the formulated processing time and the experimental result are compared for single user. The experimental system has 16 sensor nodes and a central node. Each node is a PC cluster that consists of Intel Pentium III 800MHz as CPU with 256Mbyte RAM. Each PC is general-purpose PC, which has image capturing board mounted in a PCI bus on each sensor node PC. Gigabit Ethernet connects sensor node PCs and with a central node PC. Cameras are set with 3-degree (20mm) interval on arc array and about 35cm distance from object plane. In our system for I = 160 x 120 color image. Block-based interpolation method [3] is used for arbitrary view generation. Note that in our experimental system, we have one communication between two nodes in each time. Fig. 10 shows that the theoretical analysis is close to experimental result. Therefore, the theoretical analysis is reliable to be used for designing an optimized network.

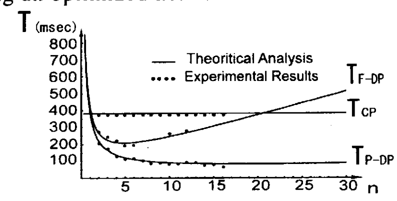
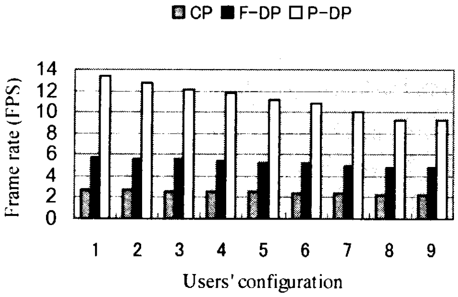


Fig. 10: Comparison of processing time of theoretical analysis with experimental system result for single user case.

In Fig. 11 we have compared network performance of our experimental system for CP and DP according to the frame and communication rates of m=10 users versus some users' configurations. In these graphs user's configurations are

(1) 10 x SU or IU: 10 x 1 user/"two views"
(2) S-DU: 5 x 2 user/"two views"
(3) S-DU: 3 x 3 user/"two views" + 1 x 1user/"two views"
(4) C-DU: 5 x 2 user/"two views"
(5) S-DU: 2 x 5 user/"two views"
(6) C-DU: 3 x 3 user/"two views" + 1 x 1user/"two views"
(7) C-DU: 2 x 4 user/"two views" + 1 x 2user/"two

views"
(8) C-DU: 2 x 5 user/"two views"
(9) C-DU: 1 x 10 user/"two views".



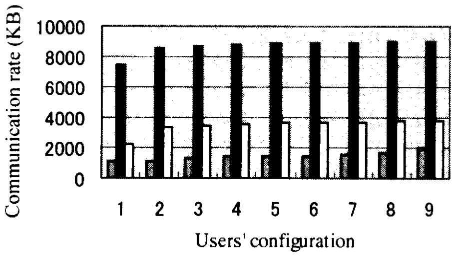


Fig. 11: Frame and Communication rates of CP and DP in network vs. user's configuration of 10 users

For "m=10" DU, using CP "n_t=1x10=10", F-DP "n_t=5x10=50" and P-DP "n_t=18x10=180" nodes are needed. Frame rate using P-DP is the best for our system. Communication rate in DP methods are higher than CP. As it can be seen, generally P-DP has optimum performance considering processing time (i.e. frame rate) and communication amount, however communication traffic and the necessary number of nodes for optimal operation of network are higher than F-DP and CP methods.

6. Conclusion

We proposed optimum network architecture to distribute the processing task in CSN to generate arbitrary view for multiuser. The network processing time is formulated, which is quite reliable to optimize the network according to experimental results. The best processing method can be chosen based on communication delay and processing ability of nodes. P-DP can optimize the network for multiuser case, considering its high speed in processing, with relatively low amount of inter-node communication.

In our future research, we will consider different view resolution of user and communication delay among nodes to optimize the network for more practical situations. In multiuser case, traffic of the network is an important factor that should be considered for better optimization result.

Furthermore, the number of necessary cameras images for arbitrary view generation is increased when the virtual camera location (i.e. user) gets distance from the camera plane. Thus, more interpolation is needed to generate the requested viewpoint out of the cameras' optical center plane. This is similar to multiuser case, however more nodes are involved to generate a viewpoint for one user. Therefore, for optimal operation there will be a trade-off among total numbers of nodes, number of users and the maximum distance of users from cameras' plane.

References

[1] P. Na Bangchang, T. Fujii, M. Tanimoto, Experimental System of free viewpoint television, Proc. of SPIE Symposium on Electronic Imaging, Santa Clara, CA, USA, Vol. 5006-66, pp. 554-563, Jan 2003.

[2] C. C. Shen, C. Srisathapornphat, C. Jaikaeo, "Sensor information networking architecture and applications", IEEE Personal Communications, Vol. 8, pp. 52 –59, Aug. 2001.

[3] A. Nakanishi, T. Fujii, T. Kimoto, M. Tanimoto, "Ray-Space Data Interpolation by Adaptive Filtering Using Locus of Corresponding Points on Epipolar Plane Image", The journal of the institute of Image information and Television Engineers (ITE), Vol. 56, No. 8, pp. 1321-1327, Aug. 2002. (In Japanese)

[4] M. P. Tehrani, T. Fujii, M. Tanimoto, "An Interpolation Method for Ray Space Data using Offset Blocks", Proc. 3D Image Conference, Tokyo, Japan, pp.129-132, July 2003.

[5] M. P. Tehrani, T. Fujii, M. Tanimoto, "Distributed Processing Method for Arbitrary View Generation in Camera Sensor Network", Proc. of SPIE Symposium on Electronic Imaging, Santa Clara, CA, USA, Vol. 5022-65, pp.554-561, Jan. 2003.

[6] Jack Y. B. Lee, "On a Unified Architecture for Video-on-Demand Service", IEEE Tran. On multimedia, Vol. 4, No. 1, pp. 38-47, March 2002.

[7] T. Fujii, T. Kimoto, and M. Tanimoto, "A New Flexible Acquisition System of Ray-Space Data for Arbitrary Objects", IEEE Trans. on Circuit and Systems for Video Technology, Vol. 10, No. 2, pp. 218-224, March 2000.

[8] F. Fuseillo, E. Trucco, and A. Verri, "A Compact Algorithm for Rectification of Stereo Pairs", Machine Vision and Applications, Vol. 12, pp. 16-22, 2000.