オーディオビジュアル複合情報処理　39－19
(2002. 12. 17)

カメラセンサーネットワークにおける多視点画像生成のための計算処理分散化

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　我々は新視点画像生成に際の処理速度向上のため、計算処理を分散化する。提案法において、我々は最小平均二乗誤差に基づいた適応フィルタ探索をセンターノードを介さずそれぞれのノード間の通信のみによって行う。また我々は従来法では中央集約化された上で行われていた新視点画像生成のための計算処理についても、FIS-DP 法と PIS-DP 法の２手法を提案し分散化する。すなわち、FIS-DP 法では２つのノード間で画像全体を共有するが、適応フィルタ探索を相補的に行い、PIS-DP 法では画像の共有を相補的に分割するかわりに、適応フィルタ検索をフルに行う。我々は両手法においてネットワーク全体の処理速度が従来の中央集約化計算手法（CP 法）を上回ることを確認した。また、PIS-DP 法は他手法（FIS-DP 法、CP 法）と比較して処理速度と通信効率の点で良好な結果を示した。PIS-DP 法を用いることによって、我々はカメラセンサーネットワーク内でのリアルタイムに自由視点画像を生成することが可能になった。

キーワード　処理分散化、 自由視点、カメラセンサーネットワーク

Decentralized Multi View Image Processing in Camera Sensor Network

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　To speed up the processing time, we decentralized the processing tasks to generate new view. In proposed methods, each sensor node processes part of corresponding search algorithm based on MSE minimization to generate the interpolated image with local communication between sensor nodes. In conventional method, all the processing tasks to generate a new view were centralized in central node. We proposed two methods, which are Fully Image Shared Decentralized Processing (FIS-DP), and Partially Image Shared Decentralized Processing (PIS-DP). In FIS-DP, full image transformation should be done between neighbor camera sensor nodes, but the correspondence check is performed in different block intervals. In PIS-DP, the correspondence check is performed in part of images in each node by sharing part of captured images. The experimental results of FIS-DP, and PIS-DP show that we can increase the processing speed of the whole network to generate a new view in compare with Centralized Processing (CP) method. The PIS-DP method is recommended according to its better processing speed and communication efficiency in compare with other methods (FIS-DP, and CP). Using PIS-DP, will let us to perform real-time free view generation in camera sensor network.

Keyword　Decentralized Processing, Multi View, Camera Sensor Network

1. Introduction

Three-dimensional (3-D) senses are generally visualized using conventional 2-D display techniques. Rendering photo-realistic 2-D views of 3-D objects at interactive frame-rate is therefore a vital aspect in 3-D video technology. 3-D rendering commonly relies on geometry description, texture information, and illumination specifications. Rendering quality, as well as frame, is determined by model complexity and computation time spent on simulating global illumination effects. Trading off rendering frame-rate versus image quality is often inevitable, and 3-D geometry acquisition from real-world objects can prove problematic if surface are not well defined (e.g. hair, fur, smoke) [5].

In this research, we address the communication efficiency and processing speed in camera sensor network. Camera sensor network [1] as a new advent of technology is a network that each sensor node can capture video signals and, process them and communicate with other nodes. In this network, many cameras have been installed in a dense configuration with overlap as sensor nodes and a central node. The arbitrary view is requested from central node or user. Interpolation to generate arbitrary view using ray space data [2,3,4] by adaptive filtering is the processing task in our network. The processing time using centralized processing (CP) method is slow. So, we two proposed decentralized (DP) methods to speed up the processing task as well as considering communication efficiency in network.

2. Arbitrary view generation

2.1. Rendering methods

The rendering methods are roughly classified into Model-Based Rendering (MBR) and Image-Based Rendering (IBR). MBR [5] describes the shapes of 3-D objects in polygons, and rendering based on the physical model generates the image. However, it is difficult to model all things in polygons, and a problem remains in real time play because of the huge calculation time. IBR is a technique of reproducing space using the real images from multi-cameras, and can be realized by comparatively light processing. "Morphing" and "Mosaicing" are categorized into this approach.

Ray-space data [2,3,4] interpolation by adaptive filtering [6] is a method to generate a new view, which has been used in this research as processing task of network. Ray space data is a common data format for 3D communication. The important feature of ray-space is that the image information seen from a certain viewpoint is given as one of a sub space of ray-space. To generate the part of ray data, which can be captured by a virtual camera (arbitrary view) between two cameras, the interpolation task should be done.

2.2. Interpolation

Interpolation method, which will be used in this research, is pixel based and correspondence check is based on Mean Square Error (*MSE*) minimization of correlated blocks [6]. So, the processing task is to up-sample the parallel epipolar line of each images after rectification [7]. Rectification determines a transformation of image plane such that pairs of conjugate epipolar lines become collinear and parallel to one of image axis. The best gradient adaptive filter (with the least *MSE*) should be selected for corresponded pixels to generate intermediate images between two images by applying the selected adaptive filter. See Fig. 1

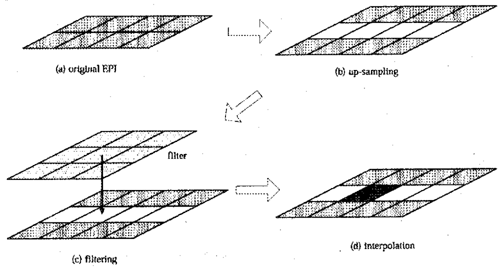

Fig. 1 Interpolation procedure [6].

Checking the minimum *MSE* value between blocks of pixels as it has been shown in Fig. 2, finds the best gradient adaptive filter to generate interpolated pixel. Mathematically, *MSE* can be shown as equation (1).

$$MSE = \sum_{i=1}^{k} (A_i - B_{k-i})^2 \qquad (1)$$

where k is number of pixels in each block and A_i and B_i shows pixels in each block.

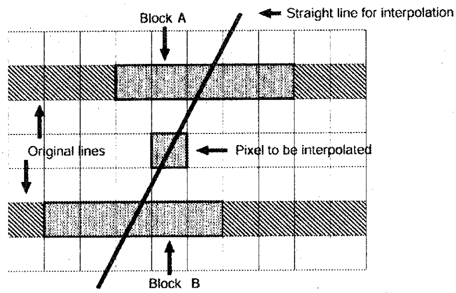

Fig. 2 Block based Corresponding search algorithm [6].

The adaptive filter for corresponded pixels is a mask, which shows the best disparity shift between two blocks of pixels of two different views to generate the interpolated pixel. This interpolation method is object independent.

3. Conventional Processing Method

3.1. Centralized Processing (CP)

In conventional method, first, all captured images data by each camera sensor node are sent to central node. Then all processing task to generate a new view (i.e.

rectification and interpolation) is done in central node. It should be mentioned that the rectification transformation matrices is calculated after calibrating cameras in network. So, the rectification process is just to apply the transform matrices to captured images. The camera calibration is done while network is initialized.

4. Proposed Decentralized Processing (DP)

The arbitrary view image between two neighbor cameras is supposed to be generated. Lets consider the network configuration in Fig. 3 with inter-node communication. To decentralize the processing task between nodes (i.e. two sensor nodes and a central node), we propose two methods as follow.

- *Fully Image Shared Decentralized Processing (FIS-DP)*
- *Partially Image Shared Decentralized Processing (PIS-DP)*

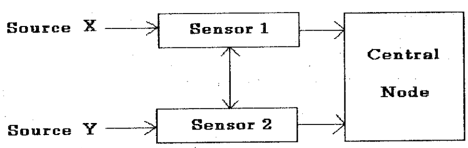

Fig. 3 Communication protocol in proposed camera sensor network

4.1. Fully Image Shared Decentralized Processing (FIS-DP)

In this method, rectified image information of camera sensor node 1 should be fully transferred to camera sensor node 2, and vise versa (fully shared). The correspondence check is performed in different block intervals of pixels as it has been shown in Fig. 4.

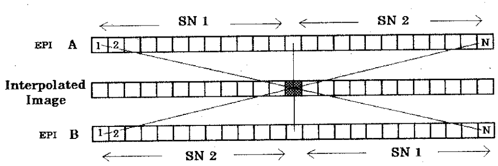

Fig. 4 Correspondence searching task in two camera sensor nodes for FIS-DP method.

Each sensor node after finding the best adaptive filter and performing the interpolation task, will transmit interpolated pixel value (8-bit) and normalized *MSE* value (8-bits) to central node. The *MSE* is normalized to *256k/2*. See equation (2).

$$\overline{MSE}=\sum_{i=1}^{k/2}(A_i - B_{\frac{k}{2}-i})^2 \Big/ (128k) \qquad (2)$$

where k is number of pixels in each block that *MSE* is measured, and A_i and B_i shows pixels in each block, and $\overline{MSE}$ is normalized value of *MSE*. The to be transmitted information to central node is shown in Fig 5, and equal to two black and white image data.

Fig. 5 Sensor node to central node Communicating frame in FIS-DP method.

Central node chooses the interpolated pixel value with the minimum $\overline{MSE}$ value between the data, which has received from each sensor node.

4.2. Partially Image Shared Decentralized Processing (PIS-DP)

In this framework, the captured images by each sensor node are vertically divided into two equal parts. We call them upper and downer parts. The upper part of captured image by sensor node 1 is rectified and transmitted to sensor node 2, and downer part of captured image by sensor node 2 is rectified and sent to sensor node 1 (i.e. partially shared). See Fig. 6.

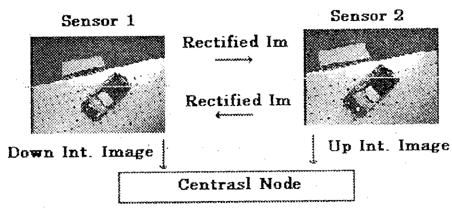

Fig. 6 PIS-DP method communications protocol.

Then, the interpolation task will be done in each sensor nodes for each part of image data. So, the data that should be transmitted to central node will be the interpolated pixels value for half of interpolated image. Central node generates the new view by adding each part to an image plane.

4.3. Performance Analysis of CP, FIS-DP, and FIS-DP Methods

In this section, amount of communication and processing speed of proposed methods and conventional method are formulated. "**I**" shows the amount of data in one black and white image. Therefore, the to be communicated data in FIS-DP, PIS-DP, and CP methods are equal to "6**I**", "2**I**", and "2**I**", respectively. The whole processing speed is equal to summation of communication delays, rectification and interpolation time. We use "$\mathbf{t_d}$"

to show communication delay to transmit an image "**I**" between two nodes, and "$\mathbf{T_p}$" for rectification and interpolation time to generate an interpolated image. So, the decentralized processing task speed for FIS-DP method, is equal to "$3\mathbf{t_d}+0.5\mathbf{T_p}$", and for PIS-DP is equal to "$\mathbf{t_d}+0.5\mathbf{T_p}$". But processing task speed for CP method is equal to "$\mathbf{t_d}+\mathbf{T_p}$".

It should be mention that the processing time "$\mathbf{T_p}$" is a lot higher than communication delay "$\mathbf{t_d}$" for our experimental system. So, FIS-DP is still faster than CP method, but not communication efficient. It can be seen that PIS-DP has the best communication efficiency and processing speed in compare with FIS-DP and CP methods.

5. Results

In experimental system, each sensor node and central 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 is mounted in a PCI bus on each sensor node PC. Gigabit Ethernet connects sensor node PCs and central node PC. Cameras are mounted randomly with 5cm interval and about 60cm distance from object plane. Our captured images are shown in Fig. 6. The communication delay for an image data (**I** = 640 x 480 Bytes) is $\mathbf{t_d}$=61.2msec. The processing time to generate an interpolated image after rectification with CP method is equal to:

$\mathbf{T_p}$ = 30.5420msec (rectification) + 952.7249msec (interpolation) = 983.27msec.

If captured images have 160 x 120 size, then we will be able to transmit 15.32 frame/sec using CP, 23.69 frame/sec using FIS-DP, and 28.94 frame/sec using PIS-DP method.

Table 1 Communication delay, Rectification and Interpolation time, Total processing time, and Transferred data for CP, FIS-DP, and PIS-DP methods. I = 640 x 480.

	CP	FIS-DP	PIS-DP
t_d (msec)	61.2	183.6	61.2
T_p (msec)	983.27	491.63	491.63
t_d+T_p (msec)	1044.47	675.23	552.83
Data (Kbytes)	614.4	1843.2	614.4

Table 1 shows the whole processing time for two proposed decentralized, and centralized processing methods with 640 x 480 image size.

The comparison of the proposed methods with Centralized Processing (CP) method shows that PIS-DP has the highest processing speed after FIS-DP, and CP has the lowest processing speed. Communication rate of CP and PIS-DP is same and better than FIS-DP. Also, PIS-DP has real-time free view video generation ability with 160x120 image size.

6. Conclusion

We proposed a network communication protocol to decentralize the processing task between nodes in camera sensor network for generating arbitrary view. PIS-DP is recommended because of its better performance than CP and FIS-DP. Also, PIS-DP method optimizes the global communication of whole network, considering its high speed in processing, even with inter-node communication. In our future research, we are going to improve our proposed methods to involve more sensor nodes to perform processing tasks and to get better image quality. Therefore, lower processing time can be obtained. Beside, we would like to gain the high correlation between nodes to have better communication efficiency, by applying an appropriate coding scheme.

References

[1] C. C. Shen, C. Srisathapornphat, C. Jaikaeo, "Sensor information networking architecture and applications", IEEE Personal Communications, Vol. 8 Issue: 4, pp. 52 –59, August 2001.

[2] T. Fujii and H. Harashima: "Data Compression and Interpolation of Multi-View Image Set" IEICE Transactions on Information and Systems Vol. E77-D No.9, pp. 987-995, 1994.

[3] T. Fujii, T. Kimoto, and M. Tanimoto, "Ray Space Coding for 3D Visual Communication", Picture Coding Symposium '96, Vol.2, pp. 447-451 (1996).

[4] 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.

[5] M. Magnor, and B. Girod, "Data Compression for Light-Filed Rendering", IEEE Trans. On Circuit and Systems for Video Technology, Vol. 10, No. 3, pp. 338-343, April 2000.

[6] 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)

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