

Fast and Robust 3D Correspondence Matching and Its Application to Volume Registration

Yuichiro TAJIMA^{†a)}, Nonmember, Kinya FUDANO^{††}, Koichi ITO[†], and Takafumi AOKI[†], Members

SUMMARY This paper presents a fast and accurate volume correspondence matching method using 3D Phase-Only Correlation (POC). The proposed method employs (i) a coarse-to-fine strategy using multi-scale volume pyramids for correspondence search and (ii) high-accuracy POC-based local block matching for finding dense volume correspondence with sub-voxel displacement accuracy. This paper also proposes its GPU implementation to achieve fast and practical computation of volume registration. Experimental evaluation shows that the proposed approach exhibits higher accuracy and lower computational cost compared with conventional method. We also demonstrate that the GPU implementation of the proposed method can align two volume data in several seconds, which is suitable for practical use in the image-guided radiation therapy.

key words: CT, MRI, registration, phase-only correlation, GPU

1. Introduction

Recently, the use of Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) data makes it possible to analyze internal organs in 3D space by reconstructing a 3D volume data from a set of slice images. In addition, the spatial resolution of medical volume data has been acceptable in the practical use of daily medical practice [1], since medical imaging technologies have been rapidly improved.

Medical volume matching is an important fundamental task for analyzing medical volume data such as volume registration, volume fusion, etc. Comparing medical volume data acquired in different days, we can observe disease progression and detect anatomical changes. Combining different types of volume data from various imaging devices, we can observe the integrated information from only one volume data. Since there is unknown transformation between medical volume data acquired in different timing, the accurate and robust medical volume matching method is required.

So far, the medical volume matching methods have been proposed, which are classified into marker-based [2], geometric structure-based [3] and voxel similarity-based [4], [5] methods. In the marker-based method, the locations of implantable makers are used to align medical volumes, while the markers have to be invasive for tomography

scans [2]. In the geometric structure-based method, ICP (Iterative Closest Point) algorithm is well-known [3]. It is difficult to align medical volume data acquired from the different imaging devices, since their structures are different from each other. The voxel similarity-based method is based on maximization of the similarity between medical volume data, and is the most popular method in medical volume matching [1]. Normalized Mutual Information (NMI) proposed by Studholme et al. [4] is widely used to evaluate similarity between medical volumes. The similarity-based method using NMI is robust in multimodality cases, while its computational cost is significantly high due to nonlinear optimization to maximize similarity. Also, the correct solutions are not obtained, if the appropriate initial values are not set.

Addressing the above problems, we propose a novel volume correspondence-based method for fast and accurate medical volume registration. The proposed 3D correspondence matching method is based on 3D Phase-Only Correlation (POC) [6]. POC is a high-accuracy image matching technique using the phase components in 2D Discrete Fourier Transforms (DFTs) of given images [7]–[10] and has been successfully applied to computer vision applications [11]–[14]. 3D POC is defined as an extended version of 2D POC. The proposed 3D correspondence matching method employs (i) a coarse-to-fine strategy using multi-scale volume pyramids for correspondence search and (ii) high-accuracy POC-based local voxel matching for finding dense volume correspondence with sub-voxel displacement accuracy.

The correspondence-based method using 3D POC is more accurate and faster than the conventional method using NMI as observed in our preliminary version of this study [6]. However, the computation time to align high-resolution volume data is still a problem even if the proposed method is used. The current and past CT data have to be aligned and compared within a practical computation time. In particular, the computation time of several seconds is suitable for practical use of volume registration in the image-guided radiation therapy. Unlike other volume registration methods, a parallel implementation can be applied to the proposed correspondence-based method, since the correspondence can be obtained for each reference point and the most of operations such as Fourier transform can be done in parallel. For the purpose of effective implementation, we propose a Graphics Processing Unit (GPU) implementation of the proposed volume correspondence matching method. The

Manuscript received June 16, 2012.

Manuscript revised October 26, 2012.

[†]The authors are with the Department of Computer and Mathematical Sciences, Graduate School of Information Sciences, Tohoku University, Sendai-shi, 980–8579 Japan.

^{††}The author is with NEC Software Tohoku, Ltd., Sendai-shi, 980–0811 Japan.

a) E-mail: tajima@aoki.ecei.tohoku.ac.jp

DOI: 10.1587/transinf.E96.D.826

GPU has been very efficient at manipulating and displaying computer graphics. Recently, the highly parallel structure of GPU makes it more effective than general-purpose CPUs for algorithms where processing of large blocks of data can be done in parallel. This effort is known as General-Purpose computation on Graphics Processing Unit (GPGPU) [15]. The GPGPU has been applied to scientific computing and video processing [16]. So far, there are some works on the GPU implementation of volume registration method with nonlinear optimization [17]. The computation of volume similarity such as NMI is not suitable for the GPU implementation, since the size of shared memory in GPU is small to compute NMI. Hence, its GPU implementation has achieved 5–7 times speedup at best [18]. Experimental evaluation using actual CT and MRI data demonstrates that the proposed method exhibits higher registration accuracy compared with conventional method, and is effective even for multimodality cases such as CT-MRI registration. Also, we demonstrate that the proposed GPU implementation of our method exhibits 10–20 times speedup compared with its CPU implementation.

The rest of the paper is organized as follows: Sect. 2 describes fundamentals of 3D POC and a set of techniques for high-accuracy volume matching using 3D POC. Section 3 describes the proposed volume correspondence matching method using 3D POC and a set of techniques for high-accuracy volume correspondence matching. Section 4 shows a proposed GPU implementation of volume correspondence matching method. Section 5 shows an application of the proposed method to rigid volume registration and its performance evaluation. Section 6 ends with some concluding remarks.

2. Fundamentals of 3D Phase-Only Correlation

This section describes fundamentals of 3D POC and techniques for high-accuracy volume matching using 3D POC.

The 3D POC function is defined as an extended version of the 2D POC function proposed in [9]. Consider two $N_1 \times N_2 \times N_3$ volume data, $f(n_1, n_2, n_3)$ and $g(n_1, n_2, n_3)$, where $n_1 = -M_1, \dots, M_1$ ($M_1 > 0$), $n_2 = -M_2, \dots, M_2$ ($M_2 > 0$), $n_3 = -M_3, \dots, M_3$ ($M_3 > 0$), and hence $N_1 = 2M_1 + 1$, $N_2 = 2M_2 + 1$, $N_3 = 2M_3 + 1$. Note that we assume here the sign symmetric index ranges $\{-M_1, \dots, M_1\}$, $\{-M_2, \dots, M_2\}$ and $\{-M_3, \dots, M_3\}$ for mathematical simplicity. The discussion could be easily generalized to non-negative index ranges with power-of-two volume size.

The 3D Discrete Fourier Transforms (3D DFTs) of $f(n_1, n_2, n_3)$ and $g(n_1, n_2, n_3)$ are given by

$$\begin{aligned} F(k_1, k_2, k_3) &= \sum_{n_1, n_2, n_3} f(n_1, n_2, n_3) W_{N_1}^{k_1 n_1} W_{N_2}^{k_2 n_2} W_{N_3}^{k_3 n_3} \\ &= A_F(k_1, k_2, k_3) e^{j\theta_F(k_1, k_2, k_3)}, \end{aligned} \quad (1)$$

$$\begin{aligned} G(k_1, k_2, k_3) &= \sum_{n_1, n_2, n_3} g(n_1, n_2, n_3) W_{N_1}^{k_1 n_1} W_{N_2}^{k_2 n_2} W_{N_3}^{k_3 n_3} \\ &= A_G(k_1, k_2, k_3) e^{j\theta_G(k_1, k_2, k_3)}, \end{aligned} \quad (2)$$

where $k_1 = -M_1, \dots, M_1$, $k_2 = -M_2, \dots, M_2$, $k_3 = -M_3, \dots, M_3$, $W_{N_1} = e^{-j\frac{2\pi}{N_1}}$, $W_{N_2} = e^{-j\frac{2\pi}{N_2}}$, $W_{N_3} = e^{-j\frac{2\pi}{N_3}}$, and the operator \sum_{n_1, n_2, n_3} denotes $\sum_{n_1=-M_1}^{M_1} \sum_{n_2=-M_2}^{M_2} \sum_{n_3=-M_3}^{M_3}$. $A_F(k_1, k_2, k_3)$ and $A_G(k_1, k_2, k_3)$ are amplitude components, and $\theta_F(k_1, k_2, k_3)$ and $\theta_G(k_1, k_2, k_3)$ denote phase components.

The normalized cross-power spectrum of $F(k_1, k_2, k_3)$ and $G(k_1, k_2, k_3)$ is given by

$$\begin{aligned} R(k_1, k_2, k_3) &= \frac{F(k_1, k_2, k_3) \overline{G(k_1, k_2, k_3)}}{|F(k_1, k_2, k_3) G(k_1, k_2, k_3)|} \\ &= e^{j(\theta_F(k_1, k_2, k_3) - \theta_G(k_1, k_2, k_3))}, \end{aligned} \quad (3)$$

where $\overline{G(k_1, k_2, k_3)}$ denotes the complex conjugate of $G(k_1, k_2, k_3)$. Note that $\theta_F(k_1, k_2, k_3) - \theta_G(k_1, k_2, k_3)$ indicates the phase difference between two volumes. The 3D POC function $r(n_1, n_2, n_3)$ is the 3D Inverse DFT (3D IDFT) of normalized cross-power spectrum and is defined as

$$\begin{aligned} r(n_1, n_2, n_3) &= \frac{1}{N_1 N_2 N_3} \sum_{k_1, k_2, k_3} R(k_1, k_2, k_3) \\ &\quad \times W_{N_1}^{-k_1 n_1} W_{N_2}^{-k_2 n_2} W_{N_3}^{-k_3 n_3}, \end{aligned} \quad (4)$$

where \sum_{k_1, k_2, k_3} denotes $\sum_{k_1=-M_1}^{M_1} \sum_{k_2=-M_2}^{M_2} \sum_{k_3=-M_3}^{M_3}$. When two volumes are similar, their POC function gives a distinct sharp peak like a delta function. The correlation value of the peak can be used as a good similarity measure for volume matching, and the location of the peak shows the translational displacement between the two volumes. Other important properties of POC are that POC-based volume matching is not influenced by shift and brightness change, and it is highly robust against noise.

Let δ_1 , δ_2 and δ_3 represent sub-voxel displacements of $f(n_1, n_2, n_3)$ to n_1 , n_2 and n_3 directions, respectively. The displaced volume can be represented as $g(n_1, n_2, n_3) = f(n_1 - \delta_1, n_2 - \delta_2, n_3 - \delta_3)$. The 3D POC function between $f(n_1, n_2, n_3)$ and $g(n_1, n_2, n_3)$ is given by

$$\begin{aligned} r(n_1, n_2, n_3) &\simeq \frac{\alpha}{N_1 N_2 N_3} \frac{\sin(\pi(n_1 + \delta_1))}{\sin(\frac{\pi}{N_1}(n_1 + \delta_1))} \\ &\quad \times \frac{\sin(\pi(n_2 + \delta_2))}{\sin(\frac{\pi}{N_2}(n_2 + \delta_2))} \frac{\sin(\pi(n_3 + \delta_3))}{\sin(\frac{\pi}{N_3}(n_3 + \delta_3))}, \end{aligned} \quad (5)$$

where $\alpha = 1$. The above equation represents the shape of the correlation peak of the 3D POC function for common volumes that are minutely displaced with each other. The peak position of the 3D POC function corresponds to the displacement between the two volumes. We can prove that the peak value α decreases without changing the function shape itself, when small noise components are added to the original volumes. Hence, we assume $\alpha \geq 1$ in practice.

Listed below are important techniques for high-accuracy sub-voxel volume matching using 3D POC.

(A) Windowing to reduce boundary effects

Due to the DFT's periodicity, a volume can be consid-

ered to “wrap around” at a boundary, and hence discontinuities occur at every boundary in 3D DFT computation. We reduce the effect of discontinuity at volume border by applying 3D window function to the input volumes $f(n_1, n_2, n_3)$ and $g(n_1, n_2, n_3)$. In this paper, we employ the 3D Hanning window defined as

$$w(n_1, n_2, n_3) = \frac{1 + \cos(\frac{\pi n_1}{M_1})}{2} \frac{1 + \cos(\frac{\pi n_2}{M_2})}{2} \frac{1 + \cos(\frac{\pi n_3}{M_3})}{2}. \quad (6)$$

(B) Spectral weighting technique to reduce aliasing and noise effects

For natural images, typically the high frequency components may have less reliability, i.e., low S/N, compared with the low frequency components. We could improve the estimation accuracy by applying a low-pass-type weighting function $H(k_1, k_2, k_3)$ to $R(k_1, k_2, k_3)$ in frequency domain and eliminating the high frequency components having low reliability. In this paper, we employ a Gaussian version of the spectral weighting function defined by

$$H(k_1, k_2, k_3) = e^{-2\pi^2\sigma^2\left\{\left(\frac{k_1}{N_1}\right)^2 + \left(\frac{k_2}{N_2}\right)^2 + \left(\frac{k_3}{N_3}\right)^2\right\}}, \quad (7)$$

where σ is constant. In this case, the 3D POC function is defined by the 3D IDFT of $H(k_1, k_2, k_3) \times R(k_1, k_2, k_3)$, and the correlation peak model corresponding to Eq. (5) is rewritten by

$$r(n_1, n_2, n_3) = \frac{1}{N_1 N_2 N_3} \sum_{k_1, k_2, k_3} H(n_1, n_2, n_3) R(k_1, k_2, k_3) \times W_{N_1}^{-k_1 n_1} W_{N_2}^{-k_2 n_2} W_{N_3}^{-k_3 n_3} \quad (8)$$

$$\simeq \frac{\alpha}{2\pi\sigma^2} e^{-((n_1+\delta_1)^2 + (n_2+\delta_2)^2 + (n_3+\delta_3)^2)/2\sigma^2}. \quad (9)$$

This function shows the correlation peak at $(-\delta_1, -\delta_2, -\delta_3)$ as well as the peak model of Eq. (5). Note that the arbitrary functions can be applied to the POC function as a spectral weighting function depending on the intended use.

(C) Function fitting for high-accuracy estimation of peak position

We use the closed-form peak model of the POC function defined by Eq. (9) directly for estimating the peak position by function fitting. By calculating the POC function, we obtain a data of $r(n_1, n_2, n_3)$ for each discrete index n . It is possible to find the location of the peak that may exist between voxels by fitting the function Eq. (9) to the calculated data array around the correlation peak, where α , δ_1 , δ_2 and δ_3 are fitting parameters. However, this method requires an iterative computation for nonlinear function fitting, resulting in significant increase in computation time. Addressing this problem, we have proposed a Peak Evaluation Formula (PEF) that directly estimates the POC peak location from actual data array of the POC function [19]. The use of PEF allows us to eliminate iterative computation in the peak estimation. This technique can be easily extended to three

dimensions. In this paper, we apply the PEF technique to estimate the peak position of the 3D POC function.

3. Volume Correspondence Matching Using 3D POC

This section describes the details of the proposed volume correspondence matching method using 3D POC and techniques for high-accuracy volume correspondence matching. The proposed method extracts small regions (small 3D blocks) from the two volume data, estimates the translational displacement between the small blocks with sub-voxel accuracy using 3D POC, and then finds the accurate corresponding point pairs between the two volume data. In order to achieve robust correspondence matching, we employ the coarse-to-fine strategy using volume pyramids to find the corresponding point pairs between the two volume data.

Consider two volume data I and J , which are taken at different timing, or are taken from different sensors. Let $\mathbf{p} = (p_1, p_2, p_3) \in \mathbb{Z}^3$ be a coordinate vector of a reference point on I , where \mathbb{Z} is the set of integers. The problem of sub-voxel correspondence matching is to find a real-number coordinate vector $\mathbf{q} = (q_1, q_2, q_3) \in \mathbb{R}^3$ on J that corresponds to the reference point \mathbf{p} on I , where \mathbb{R} is the set of real numbers. Figure 1 shows an overview of the proposed sub-voxel correspondence matching method. We describe the details of the proposed method in the following.

Step 1: For $l = 1, 2, \dots, l_{max}$, create l -th layer volume data I^l and J^l , i.e., coarser versions of $I^0 (= I)$ and $J^0 (= J)$ by reducing I^0 and J^0 by 2^{-l} . In this paper, calculate I^{l+1} as mean values of every adjacent $2 \times 2 \times 2$ voxels in I^l as follows:

$$I^{l+1}(n_1, n_2, n_3) = \frac{1}{8} \sum_{b_1, b_2, b_3} I^l(2n_1 + b_1, 2n_2 + b_2, 2n_3 + b_3), \quad (10)$$

where $b_1, b_2, b_3 \in \{0, 1\}$. J^{l+1} is also calculated in the same way.

Step 2: For the coarsest layer $l = l_{max}$, calculate the coordinate of the reference point $\mathbf{p}^{l_{max}} \in \mathbb{Z}^3$ as follows:

$$\mathbf{p}^{l_{max}} = (\lfloor 2^{-l_{max}} p_1 \rfloor, \lfloor 2^{-l_{max}} p_2 \rfloor, \lfloor 2^{-l_{max}} p_3 \rfloor), \quad (11)$$

where $\lfloor z \rfloor$ denotes the operation to round the element of z to the nearest integer towards minus infinity. We assume that $\mathbf{q}^{l_{max}} = \mathbf{p}^{l_{max}} \in \mathbb{Z}^3$ in the coarsest layer. Hence, the coordinate of the corresponding point $\mathbf{q}^{l_{max}}$ is given by

$$\mathbf{q}^{l_{max}} = (\lfloor 2^{-l_{max}} p_1 \rfloor, \lfloor 2^{-l_{max}} p_2 \rfloor, \lfloor 2^{-l_{max}} p_3 \rfloor). \quad (12)$$

Let $l = l_{max} - 1$.

Step 3: For l -th layer, calculate the coordinate of the reference point $\mathbf{p}^l \in \mathbb{Z}^3$ as follows:

$$\mathbf{p}^l = (\lfloor 2^{-l} p_1 \rfloor, \lfloor 2^{-l} p_2 \rfloor, \lfloor 2^{-l} p_3 \rfloor). \quad (13)$$

Also, calculate the initial coordinate $\mathbf{q}^l \in \mathbb{Z}^3$ of the corresponding point \mathbf{q}^l as follows:

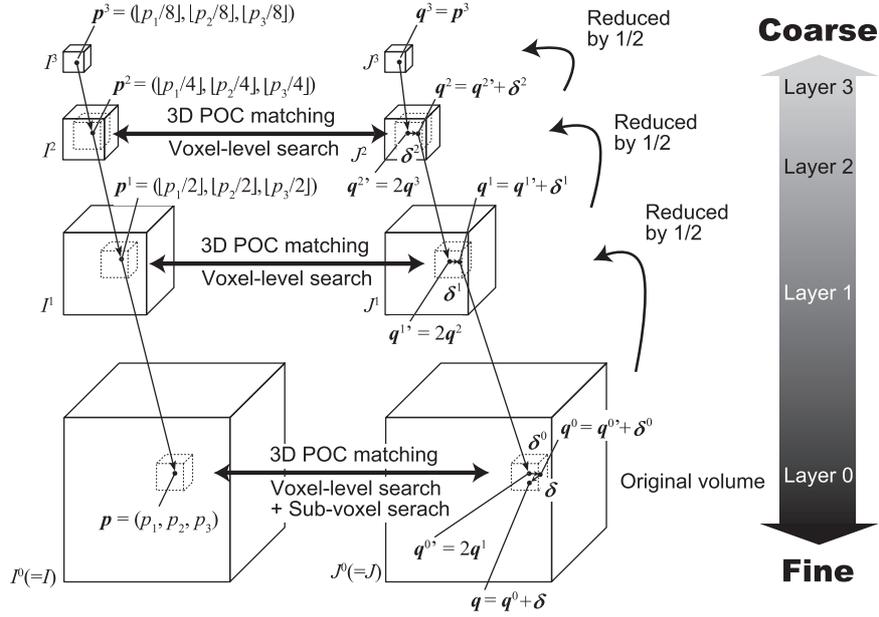


Fig. 1 Correspondence search using the coarse-to-fine strategy for $l_{max} = 3$.

$$q^{l'} = 2q^{l+1}. \quad (14)$$

Step 4: Extract two small 3D blocks f^l and g^l with their centers on p^l and $q^{l'}$, respectively, where the size of 3D blocks is $W \times W \times W$ voxels. Note that we apply zero padding to the area of volume data I^l and J^l where a portion of small 3D blocks f^l and g^l is located on outside of the volume data I^l and J^l . Estimate the displacement between f^l and g^l with voxel accuracy using 3D POC. Let the estimated displacement vector be $\delta^l \in \mathbb{Z}^3$. The l -th layer corresponding point $q^l \in \mathbb{Z}^3$ is determined as follows:

$$q^l = q^{l'} + \delta^l. \quad (15)$$

Step 5: Decrement the counter by 1 as $l = l - 1$ and repeat from Step 3 to Step 5 while $l \geq 0$.

Step 6: From the original volumes I^0 and J^0 , extract two small 3D blocks with their centers on p^0 and q^0 , respectively. Estimate the displacement between the two blocks with sub-voxel accuracy using 3D POC. Let the estimated displacement vector with sub-voxel accuracy be denoted by $\delta \in \mathbb{R}^3$. Update the corresponding points as follows:

$$q = q^0 + \delta. \quad (16)$$

Also, the peak value of the 3D POC function is obtained as a measure of reliability in local voxel matching.

Through the above procedure from Step 1 to Step 6, we find the corresponding point q on J with sub-voxel accuracy that corresponds to the reference point p on I . In the proposed method, we set many reference points on I and find their corresponding points on J to obtain dense correspondence between the medical volumes.

The followings are important techniques for high-accuracy volume correspondence matching using 3D POC.

(D) Location of Reference Points

The reference points have to be placed on the appropriate location of I , since the medical volume data have regions without any feature such as the air, which are not suitable for volume matching. In addition, in the case of the medical volume data acquired with different types of imaging devices, the reference points have to be placed on the regions having the common features between the volumes. If the reference points are placed on the regions of I having different information between the volumes, most of the obtained corresponding points on J are not accurate, so these are outliers. To address the above problems, we automatically detect the regions having the common features between the volumes and set the reference point on the appropriate location of I . For example, in the case of CT and MRI, the reference points are placed on bones and skin which can be automatically detected by using the CT value. In general, the CT values are predefined for each tissues, although we observe small variations of the CT values depending on the imaging devices and conditions. In this paper, we employ the range of the CT values $-100 \sim 0$ [HU] to detect skin and ≥ 400 [HU] to detect bones.

(E) Outlier Removal

We propose an outlier removal technique using the peak value of the POC function as a measure of correspondence reliability. The POC function between the local blocks exhibits the correlation peak depending on the similarity between the local blocks. If the local blocks are similar, the correlation peak approaches 1. When the peak value of the POC function between the local blocks is below a certain threshold, the corresponding point q is regarded as an outlier. In this paper, we employ 0.1 as the threshold value for outlier removal, where this threshold value is empirically determined. Instead of the constant threshold as men-

tioned above, we can determine the threshold using statistic approaches such as mean or standard deviation of peak values. In order to improve the robustness against outliers, we also employ RANSAC (RANdom SAmple Consensus) [20] to eliminate outliers from the corresponding point pairs obtained by the proposed method.

4. GPU Implementation

In this paper, we use the CUDA (Compute Unified Device Architecture) programming model [21] to implement the proposed correspondence matching method on the GPU. CUDA provides efficient parallel computing using both task-based and data-based parallelism.

Figure 2 shows a data parallel programming model in CUDA. The execution model in CUDA consists of two concepts: a kernel and a program. A kernel is a basic execution unit in CUDA. A program is a collection of kernels and internal functions. The program invokes a kernel over an index space called a grid. A single kernel instance at a point in the index space is called a thread. Threads are also grouped into thread blocks as shown in Fig. 2. The data-parallel execution is achieved by executing multiple thread blocks in parallel.

Figure 3 shows a memory model in CUDA. CUDA handles five memory spaces such as register, local, global, constant and texture. The global memory permits access to all threads in all thread blocks and has a large amount of capacity with long latency. The local memory is shared by all the threads in each thread block. The register memory can only be used by a thread. The constant memory may be used by all the thread blocks to store read-only data. The local and register memories can be accessed faster than the global memory, since the substance of the local and register memories is located on each core. Note that the capacity of the local and register memories is small. In the case that the amount of local memory per thread block or register memory per thread is increased, it results in reducing the number of parallel execution of threads for each core.

We present the GPU implementation techniques for the proposed volume correspondence matching method described in Sect. 3. According to the features of CUDA described in the above, we employ the following two techniques. Note that we employ the CUDA Fast Fourier Transform library (cuFFT) to calculate 3D POC function.

(i) Parallel execution for each reference point and voxel

The volume correspondence matching is to find a corresponding point on the input volume data that corresponds to the reference point in the reference volume data. The volume correspondence matching for each reference point is done in parallel, since this process is independent for each reference point. So, we assign one thread block to the correspondence matching for one reference point. In the process for each reference point, we extract the local voxel around the reference point and the search window in the input volume data, and perform the local voxel matching using 3D POC. The above process is independent for each voxel in

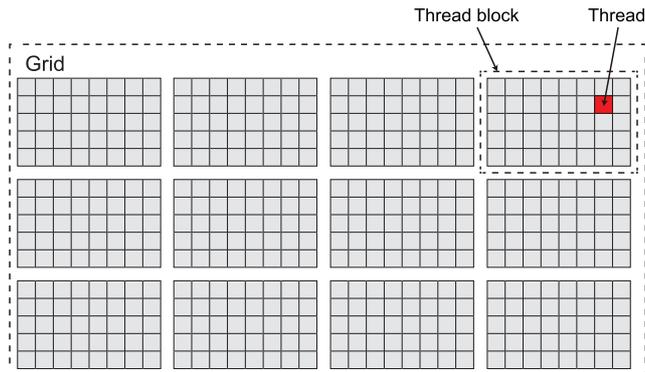


Fig. 2 Data parallel programming model in CUDA [21].

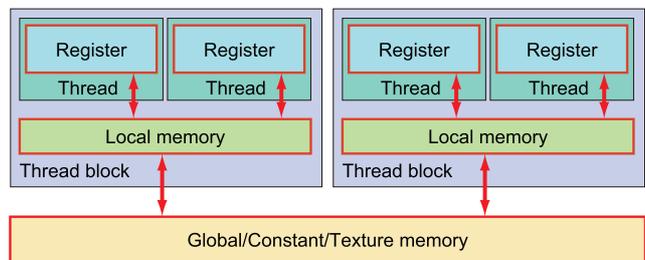


Fig. 3 Memory model in CUDA [21].

the search window. So, we assign one thread to the process for one voxel in the search window.

(ii) Size of thread block

The parallel processing is performed using $N \times N \times N$ threads per one thread block for the search window with $N \times N \times N$ voxels. Depending on the size of the search windows, a large number of resources for each thread block may be consumed and then the number of active thread blocks in each core may be reduced. Addressing this problem, we use $N \times N \times N'$ threads per one thread block and repeat $\lceil \frac{N}{N'} \rceil$ times. Note that we have to empirically determine the optimal N' , since the optimal N' depends on the GPU architectures and the size of search windows.

5. Rigid Volume Registration

We apply the proposed volume correspondence matching method to rigid volume registration. Registration between medical volume data such as CT and MRI is one of the important techniques in the field of medical image processing. The deformation between medical volume data is observed due to soft organs, imaging devices, temporal change of organs, etc. even if they are taken from the same regions of the subject. To compare volume data acquired in different days or combine different types of volume data from various imaging devices, the accurate and fast volume registration method is required. Addressing this problem, we propose a novel volume registration method using POC-based volume correspondence matching.

First, we obtain a set of accurate corresponding point

pairs between the two volumes using the proposed correspondence matching method with the techniques (D) and (E). Next, we estimate the parameters of the rigid transformation such as the rotation matrix \mathbf{R} and the translation vector \mathbf{t} using the corresponding point pairs. The parameters of the rigid transformation can be estimated from more than 3 corresponding point pairs. To estimate the parameters, we use the least-squares fitting algorithm proposed in [22], which calculates least-squares solutions of \mathbf{R} and \mathbf{t} based on the singular value decomposition.

If there are large deformation between the two volumes, it may be difficult to align the volumes by one-time procedure of correspondence matching and parameter estimation. Addressing this problem, we iteratively estimate the parameters of the rigid transformation. The reference volume is transformed by the rigid transformation with the estimated parameters so as to reduce the deformation between the volumes. We perform the correspondence matching between the transformed reference volume and the input volume and then estimate the parameters of the rigid transformation. By repeating the above procedure such as correspondence matching, parameter estimation and volume transformation, we can precisely align the two volume data even if the volume data have large deformation. In this paper, the procedure is repeated until the rotation angle between two volumes is below threshold, where the rotation angle is calculated from the rotation matrix \mathbf{R} . The threshold is 1° in this paper. We empirically confirm that the estimated parameters of the rigid transformation is converged, even if the volumes are simultaneously rotated by $\pm 30^\circ$ for roll, pitch and yaw. Unless the subject deliberately changes his/her position during two acquisition, the rotation angle between the volumes is not over $\pm 30^\circ$ for roll, pitch and yaw. Hence, we can achieve the robust rigid volume registration by iteratively estimating parameters of the rigid transformation.

6. Experiments and Discussion

We evaluate the accuracy and computation time of the rigid volume registration using actual CT and MRI data. We compare the registration accuracy of the proposed and conventional methods in Sect. 6.1, while we describe the effectiveness of the GPU implementation of the proposed method in Sect. 6.2.

6.1 Registration Accuracy

In this paper, we use the public medical volume database provided by RIRE (Retrospective Image Registration Evaluation Project) [23] to evaluate the registration accuracy of the proposed and conventional methods. The RIRE database consists of one training dataset with ground-truth registration parameters and 16 testing dataset without ground-truth registration parameters. Each dataset has one CT, three MRI (T1, T2, PD) and one PET taken from a subject. In this paper, we use CT and MRI (T1) to evaluate the registration

accuracy both for training and testing datasets.

6.1.1 Experiments with RIRE Training Dataset

In this experiment, we use one CT and one MRI data in the RIRE training dataset as shown in Fig. 4. The CT and MRI data are rescaled and zero-padded to $512 \times 512 \times 177$ isotropic voxels with spatial resolution $0.65 \times 0.65 \times 0.65$ mm.

First, we align CT and MRI data using the ground-truth registration parameters supplied by RIRE project. Then, we transform MRI data using random translation and rotation, where $N (= 8)$ groups are generated in this paper. We randomly select the translation $\hat{\mathbf{t}}_i$ and rotation $\hat{\mathbf{r}}_i$ so as to have $\|\hat{\mathbf{t}}_i\|$ within $[-20, 20]$ voxels and $\|\hat{\mathbf{r}}_i\|$ within $[-20, 20]$ degrees. Next, we estimate the translation \mathbf{t}_i and rotation \mathbf{r}_i using the rigid volume registration methods, where $i = 1, \dots, N$. Finally, the registration accuracy for translation and rotation is evaluated by Euclid distances $\|\mathbf{e}_i^t\| = \|\mathbf{t}_i - \hat{\mathbf{t}}_i\|$ and $\|\mathbf{e}_i^r\| = \|\mathbf{r}_i - \hat{\mathbf{r}}_i\|$, respectively.

In this experiment, we use the rigid volume registration method proposed by Studholme et al. [4] as a conventional method, which is one of the well-known registration methods for medical volume data. This method estimates the parameters of rigid transformation by maximizing the NMI value using the nonlinear optimization approach, where the NMI value is calculated using the overlapped area between CT and MRI data. To implement the conventional method, we use the MATLAB and C programs available in [24], where we use the version released on 15 Apr 2009. We employ the parameters of the proposed method as shown in Table 1.

Figure 5 shows the registration results using the conventional and proposed methods. As shown in Fig. 5 (c), the conventional method cannot align CT and MRI data, since the MRI data is still rotated. On the other hand, as shown in Fig. 5 (d), the proposed method correctly aligns CT and MRI data, since the position of MRI is almost the same as the true position of MRI as shown in Fig. 5 (a). The MRI data in Fig. 5 (b) is transformed by

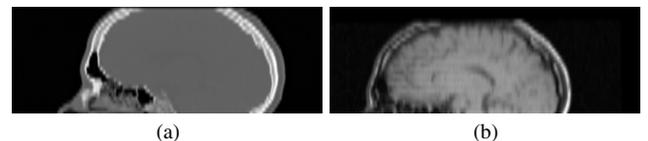


Fig. 4 Volume data in RIRE training dataset: (a) CT and (b) MRI.

Table 1 Parameters of the proposed method.

# of reference points	1,000
Block size W [voxel]	$32 \times 32 \times 32$
Spectrum weighting function	Gaussian function ($\sigma = 0.5$)
# of layers l_{max}	4

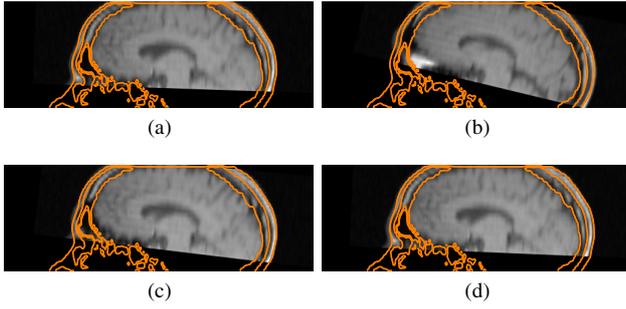


Fig. 5 Registration results for RIRE training dataset: (a) CT and MRI data aligned by the ground-truth registration, (b) CT and MRI data after manual transformation of MRI data, (c) registration result using the conventional method and (d) registration result using the proposed method, where the orange lines indicate edges extracted from the CT data.

$$\mathbf{R}_{\text{manual}} = \begin{bmatrix} 0.9915 & 0.1094 & -0.0701 \\ -0.0932 & 0.9748 & 0.2029 \\ 0.0906 & -0.1946 & 0.9767 \end{bmatrix},$$

$$\mathbf{t}_{\text{manual}} = \begin{bmatrix} -3.6839 \\ 7.9808 \\ 8.5750 \end{bmatrix},$$

where the unit of $\mathbf{t}_{\text{manual}}$ is voxel. The estimated parameters using the NMI-based and proposed methods are

$$\mathbf{R}_{\text{NMI}} = \begin{bmatrix} 0.9930 & 0.1079 & -0.0490 \\ -0.1017 & 0.9883 & 0.1139 \\ 0.0607 & -0.1081 & 0.9923 \end{bmatrix},$$

$$\mathbf{t}_{\text{NMI}} = \begin{bmatrix} -2.8218 \\ 3.0131 \\ 5.9425 \end{bmatrix},$$

$$\mathbf{R}_{\text{POC}} = \begin{bmatrix} 0.9915 & 0.1094 & -0.0708 \\ -0.0935 & 0.9757 & 0.1982 \\ 0.0907 & -0.1899 & 0.9776 \end{bmatrix},$$

$$\mathbf{t}_{\text{POC}} = \begin{bmatrix} -2.9968 \\ 9.9321 \\ 9.3035 \end{bmatrix},$$

respectively. As a result, the parameters \mathbf{R}_{POC} and \mathbf{t}_{POC} estimated by the proposed method are close to the ground truth $\mathbf{R}_{\text{manual}}$ and $\mathbf{t}_{\text{manual}}$ compared with the parameters \mathbf{R}_{NMI} and \mathbf{t}_{NMI} estimated by the NMI-based method.

Figure 6 shows the registration errors for the conventional and proposed methods, where the horizontal and vertical axes indicate rotation and translation errors, respectively. The average translation and rotation errors of the conventional method are 2.02 voxels and 1.45 degrees, respectively, while those of the proposed methods are 1.44 voxels and 0.45 degrees, respectively. As a result, the registration accuracy of the proposed method is higher than that of the conventional method.

The computation time of each method is evaluated on the system as shown in Table 2. The computation time is about 74 seconds for the proposed method and about 1,567 seconds for the conventional method. Thus, the proposed method not only exhibits the accurate rigid volume registra-

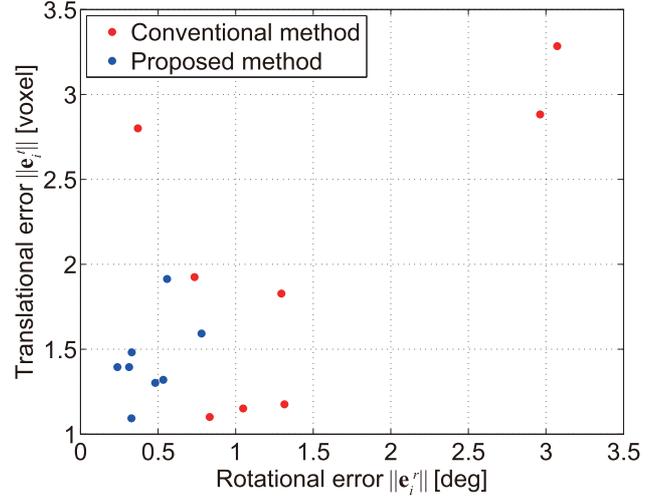


Fig. 6 Registration errors for RIRE training dataset.

Table 2 System specifications.

Item	Specification
CPU / Memory	Intel Core i3-530 2.93 GHz / 4 GB
OS	Windows 7 Pro 64 bit
GPU	NVIDIA GeForce GTX 580
CPU Implementation	MATLAB 7.13.0 / Microsoft VC++ 6.0
GPU Implementation	Microsoft VC++ 6.0 / CUDA Toolkit 4.0

tion but also achieves about 21-time faster than the conventional method.

6.1.2 Experiments with RIRE Testing Dataset

In this experiment, we use CT and MRI data taken from 16 subjects available in the RIRE testing dataset. The size of CT and MRI data is different from each subject. In order to have the same voxel size each other, the CT and MRI data are rescaled and zero-padded to $256 \times 256 \times 98$ isotropic voxels with spatial resolution $0.8 \times 0.8 \times 0.8 \text{ mm} \sim 1.3 \times 1.3 \times 1.3 \text{ mm}$.

Since this dataset does not have the ground-truth registration parameters, we use the linearity between the given and estimated transformation to evaluate the registration accuracy of the methods. We transform MRI data using random translation and rotation, where $N (= 8)$ groups are generated in this paper. We randomly select the translation $\hat{\mathbf{t}}_i$ and rotation $\hat{\mathbf{r}}_i$ so as to have $\|\hat{\mathbf{t}}_i\|$ within $[-10, 10]$ voxels and $\|\hat{\mathbf{r}}_i\|$ within $[-10, 10]$ degrees. Then, we estimate the translation \mathbf{t}_i and rotation \mathbf{r}_i using the rigid volume registration methods, where $i = 1, \dots, N$. Let \mathbf{r}_o and \mathbf{t}_o be the unknown offsets between CT and MRI data. The problem considered in this experiment is to estimate $\hat{\mathbf{r}}_i + \mathbf{r}_o$ and $\hat{\mathbf{t}}_i + \mathbf{t}_o$. If the CT and MRI data are correctly aligned, the estimated parameters are $\mathbf{r}_i = \hat{\mathbf{r}}_i + \mathbf{r}_o$ and $\mathbf{t}_i = \hat{\mathbf{t}}_i + \mathbf{t}_o$. Hence, the errors $\mathbf{r}_i - \hat{\mathbf{r}}_i$ and $\mathbf{t}_i - \hat{\mathbf{t}}_i$ should be constant independent of i , since $\mathbf{r}_i - \hat{\mathbf{r}}_i = \mathbf{r}_o$ and $\mathbf{t}_i - \hat{\mathbf{t}}_i = \mathbf{t}_o$. Thus, we evaluate the registration accuracy by the standard deviations of $\mathbf{e}_i^t = \mathbf{t}_i - \hat{\mathbf{t}}_i$ and $\mathbf{e}_i^r = \mathbf{r}_i - \hat{\mathbf{r}}_i$. The standard deviation of \mathbf{e}_i^r , i.e., the rotation

error SD^r , is calculated by

$$SD^r = \sqrt{\frac{1}{N} \sum_{i=1}^N \|e_i^r - \bar{e}^r\|^2}, \tag{17}$$

where $\bar{e}^r = (1/N) \sum_{i=1}^N e_i^r$. The standard deviation of e_i^r , i.e., the translation error SD^t , is calculated in the same way.

Figure 7 shows the registration errors for the conventional and proposed methods, where the horizontal and vertical axes indicate rotation error SD^r and translation error SD^t , respectively. The average translation and rotation errors of the conventional method are 2.50 voxels and 2.77 degrees, respectively, while those of the proposed methods are 1.55 voxels and 0.44 degrees, respectively. As a result, the registration accuracy of the proposed method is higher than that of the conventional method. The computation time is about 27 seconds for the proposed method and about 432 seconds for the conventional method. Thus, the proposed method achieves about 16-time faster than the conventional method.

6.2 Computation Time

We evaluate the computation time of the proposed method with CPU and GPU implementations. In this experiment, we use the CT and MRI data as shown in Fig. 8 (a) where the size of each volume data is $512 \times 512 \times 512$ voxels. To evaluate the computation time, we use the system as shown

in Table 2. As for the CPU implementation, we use MATLAB 7.13.0 (64 bit). We employ the parameters of the proposed method as shown in Table 3. For the purpose of evaluating the computation time, the $10 \times 10 \times 10$ or $16 \times 16 \times 16$ or $32 \times 32 \times 32$ reference points are placed in a reticular pattern. Note that, as for the GPU implementation, the parameter estimation of rigid transformation and volume transformation are also implemented on GPU.

Table 4 shows the summary of the computation time of CPU and GPU implementations. Note that ‘‘Iterations’’ in Table 4 indicates the number of iterations when the iterative procedure described in Sect.5 is finished. Since the most of computation is the correspondence matching, the proposed GPU implementation contributes to speedup the volume registration. The total computation time of the

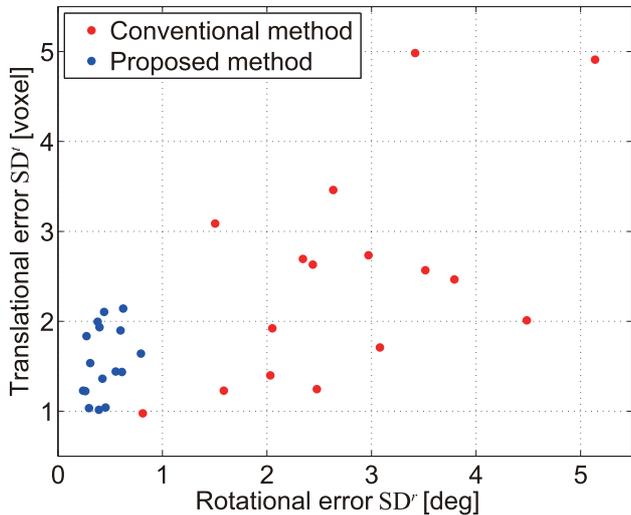


Fig. 7 Registration errors for RIRE testing dataset.

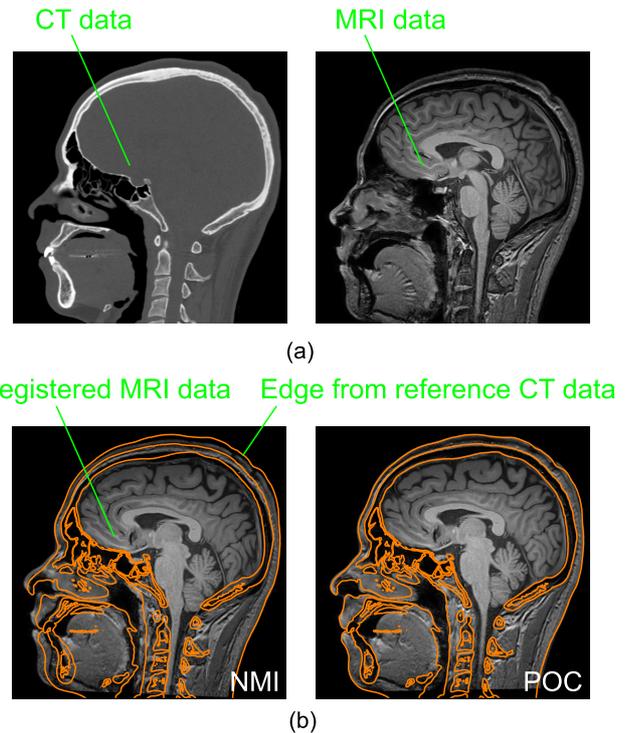


Fig. 8 Registration results of CT and MRI data: (a) before registration and (b) after registration using the proposed method.

Table 3 Parameters for computation time evaluation.

# of reference points	10^3 or 16^3 or 32^3
Block size W [voxel]	$32 \times 32 \times 32$
Spectrum weighting function	Gaussian function ($\sigma = 0.5$)
# of layers l_{max}	5

Table 4 Computation time.

Number of Points	$10 \times 10 \times 10$		$16 \times 16 \times 16$		$32 \times 32 \times 32$	
	CPU	GPU	CPU	GPU	CPU	GPU
Iterations	2	2	2	2	2	2
Correspondence Matching [sec.]	18.99	1.21	71.92	2.96	559.32	19.36
Parameter Estimation [sec.]	0.14	0.79	0.22	0.72	0.52	0.69
Transformation [sec.]	18.65	1.11	18.85	1.12	18.27	1.10
Total [sec.]	44.25	3.11	97.46	4.81	584.59	21.15

CPU implementation is 1~10 minutes, while that of the GPU implementation is 3~20 seconds. As a result, the effective GPU implementation makes it possible to achieve significant speedup of the volume registration. In practical use of the proposed method, it is only necessary to place about 1,000 reference points on the volume data. Hence, we can align two volume data in several seconds even for multimodality cases. Note that the CPU implementation of the conventional method using NMI takes 8,505 seconds to align the CT and MRI data.

7. Conclusion

This paper has proposed a novel volume correspondence matching method using 3D Phase-Only Correlation (POC). The proposed method has achieved accurate and fast correspondence matching by using (i) a coarse-to-fine strategy using multi-scale volume pyramids for correspondence search and (ii) high-accuracy POC-based local block matching. We have also proposed a GPU implementation of the proposed volume correspondence matching. Experimental evaluation using CT and MRI data has demonstrated that the proposed approach exhibits higher accuracy and lower computational cost compared with conventional method. We have also demonstrated that the GPU implementation of the proposed method can align two volume data in several seconds, which is suitable for practical use in the image-guided radiation therapy.

References

- [1] I.N. Bankman, *Handbook of Medical Image Processing and Analysis*, 2nd ed., Elsevier, 2009.
- [2] C.R.M. Jr., J.M. Fitzpatrick, M.Y. Wang, R.L.G. Jr., R.J. Maciunas, and G.S. Allen, "Registration of head volume images using implantable fiducial markers," *IEEE Trans. Med. Imaging*, vol.16, no.4, pp.447-462, Aug. 1997.
- [3] P.J. Besl and N.D. McKay, "A method for registration of 3-D shapes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.14, no.2, pp.239-256, Feb. 1992.
- [4] C. Studholme, D.L.G. Hill, and D.J. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment," *Pattern Recognit.*, vol.32, no.1, pp.71-86, Jan. 1999.
- [5] D. Rueckert, L.I. Sonoda, C. Hayes, D.L.G. Hill, M.O. Leach, and D.J. Hawkes, "Nonrigid registration using free-form deformation: Application to breast MR images," *IEEE Trans. Med. Imaging*, vol.18, no.8, pp.712-721, Aug. 1999.
- [6] K. Miyazawa, Y. Tajima, K. Ito, T. Aoki, A. Katsumata, and K. Kobayashi, "A novel approach for volume registration using 3D phase-only correlation," *Proc. Radiological Society of North America*, p.1070, Nov. 2009.
- [7] C.D. Kuglin and D.C. Hines, "The phase correlation image alignment method," *Proc. Int'l Conf. Cybernetics and Society*, pp.163-165, 1975.
- [8] Q. Chen, M. DeFRise, and F. Deconinck, "Symmetric phase-only matched filtering of Fourier-Mellin transforms for image registration and recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.16, no.12, pp.1156-1168, Dec. 1994.
- [9] K. Takita, T. Aoki, Y. Sasaki, T. Higuchi, and K. Kobayashi, "High-accuracy subpixel image registration based on phase-only correlation," *IEICE Trans. Fundamentals*, vol.E86-A, no.8, pp.1925-1934, Aug. 2003.

- [10] K. Takita, M.A. Muquit, T. Aoki, and T. Higuchi, "A sub-pixel correspondence search technique for computer vision applications," *IEICE Trans. Fundamentals*, vol.E87-A, no.8, pp.1913-1923, Aug. 2004.
- [11] K. Ito, H. Nakajima, K. Kobayashi, T. Aoki, and T. Higuchi, "A fingerprint matching algorithm using phase-only correlation," *IEICE Trans. Fundamentals*, vol.E87-A, no.3, pp.682-691, March 2004.
- [12] T. Shibahara, T. Aoki, H. Nakajima, and K. Kobayashi, "A sub-pixel stereo correspondence technique based on 1D phase-only correlation," *Proc. 2007 IEEE Int'l Conf. Image Processing*, pp.V-221-V-224, Sept. 2007.
- [13] K. Miyazawa, K. Ito, T. Aoki, K. Kobayashi, and H. Nakajima, "An effective approach for iris recognition using phase-based image matching," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.30, no.10, pp.1741-1756, Oct. 2008.
- [14] S. Sakai, K. Ito, T. Aoki, and H. Unten, "Accurate and dense wide-baseline stereo matching using SW-POC," *Proc. Asian Conf. Pattern Recognition*, pp.335-339, Nov. 2011.
- [15] J. Owens, D. Luebke, N. Govindaraju, M. Harris, J. Krüger, A. Lefohn, and T. Purcell, "A survey of general-purpose computation on graphics hardware," *Computer Graphics Forum*, vol.26, no.1, pp.80-113, 2007.
- [16] "GPGPU.org," <http://gpgpu.org>.
- [17] R. Shams, P. Sadeghi, R. Kennedy, and R. Hartley, "A survey of medical image registration on multicore and the GPU," *IEEE Signal Process. Mag.*, vol.27, no.2, pp.50-60, 2010.
- [18] F. Jung and S. Wesarg, "3D registration based on normalized mutual information: Performance of CPU vs. GPU implementation," *Proc. Workshop on Bildverarbeitung für die Medizin*, vol.5, 2010.
- [19] S. Nagashima, T. Aoki, T. Higuchi, and K. Kobayashi, "A sub-pixel image matching technique using phase-only correlation," *Proc. IEEE 2006 Int'l Symp. Intelligent Signal Processing and Communication Systems*, pp.701-704, Dec. 2006.
- [20] M.A. Fischler and R.C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Comm. of ACM*, vol.24, no.6, pp.381-395, June 1981.
- [21] D. Kirk and W. Hwu, *Programming Massively Parallel Processors*, Elsevier, 2010.
- [22] K.S. Arun, T.S. Huang, and S.D. Blostein, "Least-squares fitting of two 3-D point sets," *IEEE Trans. Pattern Anal. Machine Intell.*, vol.PAMI-9, no.5, pp.698-700, Sept. 1987.
- [23] The Retrospective Image Registration Evaluation Project. <http://www.insight-journal.org/rire/>
- [24] D. Kroon, "B-spline Grid, Image and Point based Registration - File Exchange - MATLAB Central." <http://www.mathworks.com/matlabcentral/fileexchange/20057/>



Yuichiro Tajima received the B.E. degree in information engineering, and the M.S. degree in information sciences from Tohoku University, Sendai, Japan, in 2009 and 2011, respectively. He is currently working toward the Ph.D. degree of the Graduate School of Information Sciences at Tohoku University. His research interest includes signal and image processing, and computer vision.



Kinya Fudano received the B.S. and M.S. degree in physics from Tohoku University, Sendai, Japan, in 1991 and 1993, respectively. He is currently an assistant manager of NEC Software Tohoku, Ltd. His research interest includes image processing and high-performance computational analysis.



Koichi Ito received the B.E. degree in electronic engineering, and the M.S. and Ph.D. degree in information sciences from Tohoku University, Sendai, Japan, in 2000, 2002 and 2005, respectively. He is currently an Assistant Professor of the Graduate School of Information Sciences at Tohoku University. From 2004 to 2005, he was a Research Fellow of the Japan Society for the Promotion of Science. His research interest includes signal and image processing, and biometric authentication.



Takafumi Aoki received the B.E., M.E., and D.E. degrees in electronic engineering from Tohoku University, Sendai, Japan, in 1988, 1990, and 1992, respectively. He is currently a Professor of the Graduate School of Information Sciences at Tohoku University. For 1997–1999, he also joined the PRESTO project, Japan Science and Technology Corporation (JST). His research interests include theoretical aspects of computation, VLSI computing structures for signal and image processing, multiple-valued logic, and

biomolecular computing. Dr. Aoki received the Outstanding Paper Award at the 1990, 2000, 2001 and 2006 IEEE International Symposiums on Multiple-Valued Logic, the Outstanding Transactions Paper Award from the Institute of Electronics, Information and Communication Engineers (IEICE) of Japan in 1989 and 1997, the IEE Ambrose Fleming Premium Award in 1994, the IEICE Inose Award in 1997, the IEE Mountbatten Premium Award in 1999, the Best Paper Award at the 1999 IEEE International Symposium on Intelligent Signal Processing and Communication Systems, the IP Award at the 7th LSI IP Design Award in 2005, and the Best Paper Award at the 14th Workshop on Synthesis And System Integration of Mixed Information technologies.