33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Feasibility Evaluation of a Motion Detection System with Face Images for Stereotactic Radiosurgery

Takuya Yamakawa, *Student Member, IEEE*, Koichi Ogawa, *Member, IEEE*, Hitoshi Iyatomi, *Member, IEEE,* and Etsuo Kunieda *Non Member*

Abstract— **In stereotactic radiosurgery we can irradiate a targeted volume precisely with a narrow high-energy x-ray beam, and thus the motion of a targeted area may cause side effects to normal organs. This paper describes our motion detection system with three USB cameras. To reduce the effect of change in illuminance in a tracking area we used an infrared light and USB cameras that were sensitive to the infrared light. The motion detection of a patient was performed by tracking his/her ears and nose with three USB cameras, where pattern matching between a predefined template image for each view and acquired images was done by an exhaustive search method with a general-purpose computing on a graphics processing unit (GPGPU). The results of the experiments showed that the measurement accuracy of our system was less than 0.7 mm, amounting to less than half of that of our previous system.**

I. Introduction

In stereotactic radiosurgery targeted tumors are irradiated with narrow high-energy x-ray beams, and thus any unexpected patient motion may lead to undesirable irradiation. Several methods for tracking a patient's or tumor's motion have been proposed especially with regard to abdominal breathing motion, e.g., a method with infrared markers and a stereo infrared camera[1], method with a fluoroscopic imaging system[2], method with electromagnetic transponders[3], and others[4]. In these tracking operations we can select one of two strategies: (1) turn the beam off when a small movement is detected, or (2) track the targeted organs with the beam. In our previous study we proposed a real-time contactless motion detection system [5], in which we detected the motion of the patient during the radiotherapy by tracking some areas of the patient's body with three USB (universal serial bus) cameras. And the motion tracking was performed with an active search algorithm. However, there was a limitation in the accuracy of the detected movement. This was mainly due to the resolution of acquired images (QVGA: 320x240). And if the resolution of the acquired image increased, the processing time was prolonged. As a result, the frame rate was insufficient for clinical application. On the other hand, a general-purpose computing on graphics processing unit (GPGPU) was applicable to calculate this kind of task. This prompted us to newly develop a system with this architecture. The paper describes a real-time contactless motion detection system with three USB cameras and the GPGPU.

Manuscript received March 20, 2011, revised June 18, 2011.

T. Yamakawa, K. Ogawa, H. Iyatomi are with the Department of Applied Informatics, Faculty of Science and Engineering, Hosei University, Kajinocho, Koganei, Tokyo 184-8584, Japan (corresponding author: Koichi Ogawa, Ph.D. phone & fax: +81-42-387-6189; e-mail: ogawa@ hosei.ac.jp).

E. Kunieda is with the Department of Radiology, School of Medicine, Tokai University, Isehara, Kanagawa, Japan

.

II. Materials and Method

A. Hardware configuration

In our motion detection system we used three USB cameras located around the patient's head. Fig. 1 illustrates the location of the three USB cameras and monitoring areas. We assumed that the targeted area in the radiotherapy was in a brain. Our system detected the motion of some body-parts of the patient such as ears and nose. To avoid the radiation damage of USB cameras we created as large as possible a space between the irradiation area of the target and these cameras. The movement of the patient's head was monitored with the three cameras, and the acquired images were evaluated with the template images acquired in the initial position of the patient's head.

Our system reduced the influence of the illuminance variation due to the rotation of the linear accelerator (linac) gantry by using USB cameras (DC-NCR20U, Hanwha Japan) that were sensitive to the infrared light. And we illuminated the nose and two ears with the infrared light as shown in Fig. 2. These cameras were also sensitive to the visible light, and so we cut off the visible light with a filter (IR filter, Fujifilm, Japan), and acquired images of the infrared light. We acquired images with a resolution of 640x480 (VGA) with the above three USB cameras connected to a PC (Precision M6500, [Core i7 920 Extreme edition 2GHz, 4 cores, 16 GB memory], Dell USA), and the movement of the objects was calculated with a GPGPU [Quadro FX3800M, number of stream processors: 128, 1 GB memory] to realize a high speed image processing.

B. Software configuration

Our method detects some body-parts of the patient as marker-objects in a pattern matching manner. First, we created a standard template for the specified object such as an ear or a nose being trimmed from an acquired image (Fig. 3).

For each body-part, motion tracking was performed to acquired images during the radiotherapy with each standard template mentioned above. The most likely area between the template image and an acquired image for each view was searched for with an exhaustive search method with the

using the compute unified device architecture (CUDA) [6] distributed by NVIDIA, USA. The degree of matching was evaluated with a mean squared error (MSE). Fig. 4 shows the calculation of the MSE between the template image and targeted area on a temporal frame using the texture memory of the GPU. The input image was divided into many block images with a size of 16x16, and one of the pixels in a block image was assigned to a thread of the SP. That is, calculation of the block image was done with an MP consisting of eight SPs that were calculated by 256 (=16x16) threads simultaneously. CUDA works with a unit of 32 threads, so it is efficient to select the size of a block with the multiplication of 32. To achieve a high speed matching we sampled only 1/16 of the pixels in the template image and evaluated the MSE. Fig.5 shows the code of an MSE calculation with the texture memory. The extent of movement was calculated in mm converted from pixels, and displayed on a monitor. That is, we quantified the relationship between an object-detector distance and the size of a pixel in an acquired image beforehand. In our developed graphical user interface the extents of movement of three template images were displayed graphically with three temporal images in real-time. The developed graphical user interface was shown in Fig. 6.

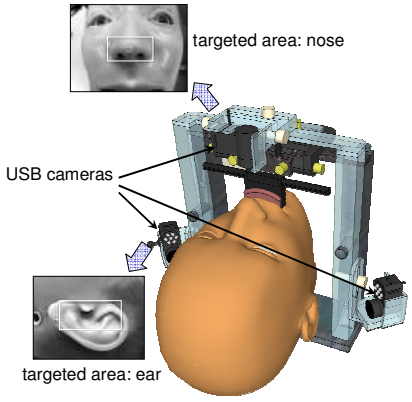

Fig. 1 Motion detection system with three USB cameras. One USB camera acquires images of the nose and the other two USB cameras acquire images of both ears.

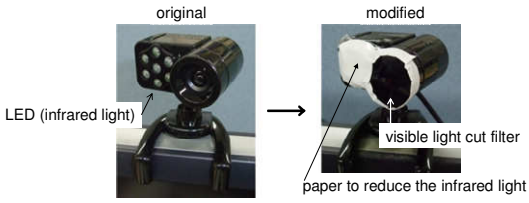

Fig. 2 USB camera and infrared light source (left). The intensity of infrared light was controlled with a piece of paper, and the visible light was cut off with a visible light cut filter in front of the camera lens.

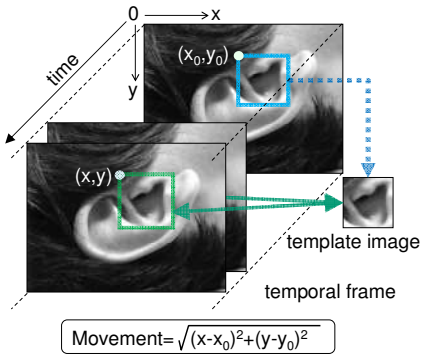

Fig. 3 Scheme of motion detection. An area of a current frame was evaluated with the template image that was defined in the initial image. The mean squared error was used for the measure.

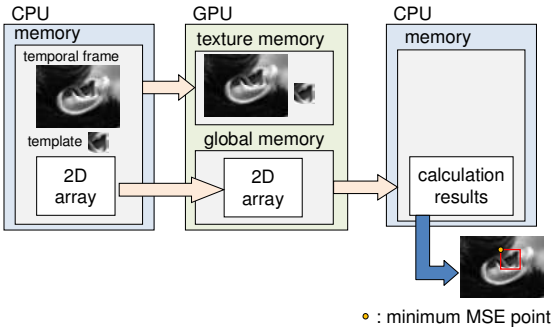

Fig. 4 Calculation of MSE with the GPU. The image data were copied from the CPU memory to the texture memory of the GPU. The results of calculation (MSE) were copied from the global memory of the GPU to the CPU memory.

C. *Experiment*

To validate the performance of our proposed system we conducted some experiments under the same conditions as those in daily stereotactic radiation therapy. Fig. 7 shows our radiation therapy room with a linac. In the room two fluorescent lamps were on the ceiling. The robustness of our motion detection system to the illuminance change of the targeted areas

(nose and both ears) was confirmed as follows, that is, we first measured the illuminance of the targeted areas with an illuminance meter by rotating the gantry of the linac. During the rotational movement of the linac gantry the shadow of the gantry appeared on a volunteer's head. In this experiment we rotated the gantry from 0 deg. to 180 deg. in 10 degree-steps, and measured the illuminance of the targeted areas. In addition, we measured the extent of the movement by averaging the results of 50 frames (5.9 seconds). The movement of the gantry was shown in Fig. 8. To compare the results with those with visible light we measured the movement of the same area without the IR filter or infrared light.

Next, we confirmed the tracking accuracy of the system under the rotational movement of the gantry. In this experiment we asked the volunteer to move his head freely during the acquisition of images for about 6 seconds (50 frames). The angular range of the gantry rotation was 50 to 90 degrees in 6 sec. In this experiment we measured the

movement of the same area without the IR filter or infrared light. To measure the true movement we used three markers and put them on the volunteer's head. We measured the extent of the movement by a manual tracing for each frame using a marker, and compared them with the output of the system that was calculated automatically.

```
// texture declaration
// an input image
texture< uchar1, 2, cudaReadModeElementType > imgTex;
// the template image :size (maskH x maskW)
texture< uchar1, 2, cudaReadModeElementType > refTex;
__global__ void kernel( float *error, int areaW, int areaH, int
maskW, int maskH ){
// get thread coordinates
int x = blockDim.x * blockIdx.x + threadIdx.x;
int y = blockDim.y * blockIdx.y + threadIdx.y;
float err = 0.0;
// in the search area
if( x < areaW && y < areaH ){
  // search area: "4" means a skip size
   for( int i = 0 ; i < maskH ; i = i + 4 ){
     for( int j = 0 ; j < maskW ; j = i + 4 ){
          // assign the position in the texture memory
          uchar1 p1 = tex2D( imgTex, x + j, y + i );
          uchar1 p2 = tex2D( refTex, j, i);
          // calculation of the squared error
          err += ( p1.x - p2.x ) * ( p1.x - p2.x );
     }
   }
   err = err / ( maskW * maskH / 16 );
 // store the calculation result for each position
  error[ areaW * y + x ] = err;
 }
}
```

Fig. 5 Codes for calculating the MSE using the texture memory of the GPU.

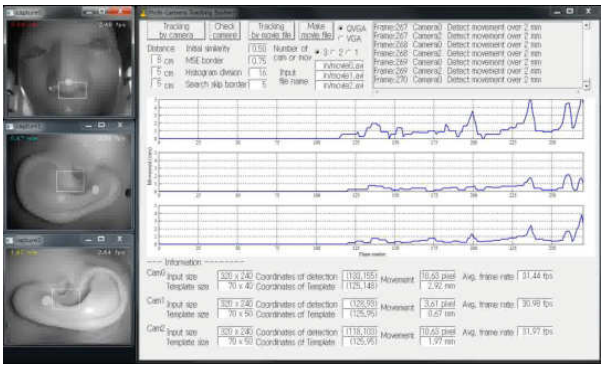

Fig. 6 Graphical user interface. The extent of the movement was displayed graphically in real time for three templates. If the amount exceeds 2 mm, the system sounds the alarm.

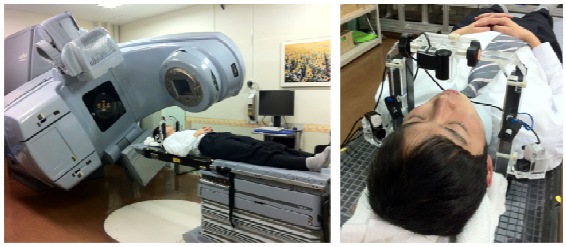

Fig. 7 Linac system and our motion detection system (left) and a close up photograph of our system with three USB infrared cameras (right).

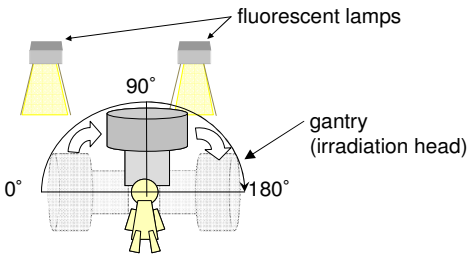

Fig. 8 Lighting condition of the radiation therapy room. The linac gantry rotates around the patient's head by 180 degrees.

III. Results

Fig. 9 shows the relationship between the variation of illuminance of the right ear and nose. Depending on the gantry angle the illuminance changed from 20 to 270 lux. And in Fig. 9 the error of detected positions is shown (left scale). The error caused by the illuminance of the visible light was significant under the dark condition, whereas the error became almost zero under the infrared lighting. Figs. 10 and 11 show the error between the output of the system and the results of the manual detection with a marker. Fig. 10 shows the results of the right ear and Fig. 11 shows those of the nose. In this experiment we asked the volunteer to move his head freely, and measured the data under the rotational movement of the gantry. As a result, the motion detection with the visible light was sometimes failed. On the other hand, the results of the motion detection with the infrared light well agreed with those of the manual tracking. In this experiment we compared the results with and without the infrared light. The results showed that our motion detection system with the infrared light accomplished good results within an average error of 0.24mm (ear) and 0.68 mm (nose). On the other hand, the measurement without the infrared light generated large errors.

IV. Discussion

For the variation of the illuminance condition the measured minimum illuminance was 20 lux and the maximum 270 lux according to the geometrical relationship between the linac gantry and fluorescent lamps. Our proposed system with the infrared light was insensitive to any changes in illuminance and tracked the target object within an error of 0.7 mm with an image size of 640x480 (VGA). In clinical situations, the location error of an x-ray beam in an irradiation should be suppressed to within 2 mm, and so the system is thought to be applicable to actual therapy.

For the calculation speed the template matching was realized with a frame rate of about 8.5, that is, the time needed for calculation was 0.117 sec. for three frames. This means that the calculation time of a frame for each camera was 0.0392 sec. On the other hand, the calculation speed with the CPU was 0.17 fps. This means the time needed for calculation was 5.88 sec for three cameras. Therefore we realized the calculation speed of approximately 50 times faster than that with the CPU. We think that a frame rate of 5 is sufficient to avoid undesired irradiation of an x-ray beam due to sudden

patient movement. From this point of view, the system is applicable for therapy with a sufficient response time. We are now improving the system GUI for actual radiation therapy.

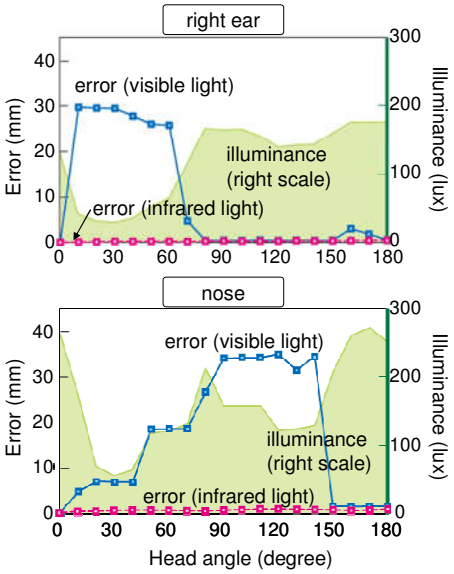

Fig. 9 Illuminance change of the targeted areas. The linac gantry rotates around the patient's head by 180 degrees. The measurement error of a targeted object with the visible light or infrared light is shown for each angle position.

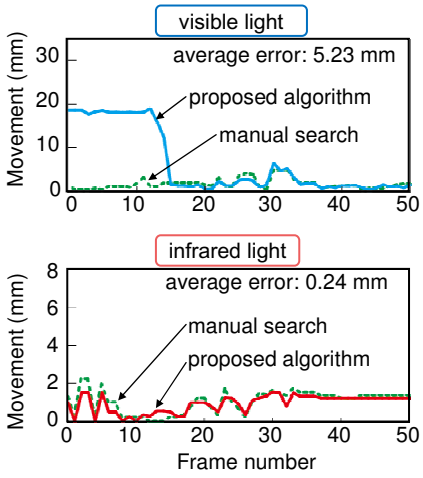

Fig. 10 Tracking accuracy of our system with the visible light or infrared light. The target object was the right ear. The system outputs are compared with the results of manual tracking.

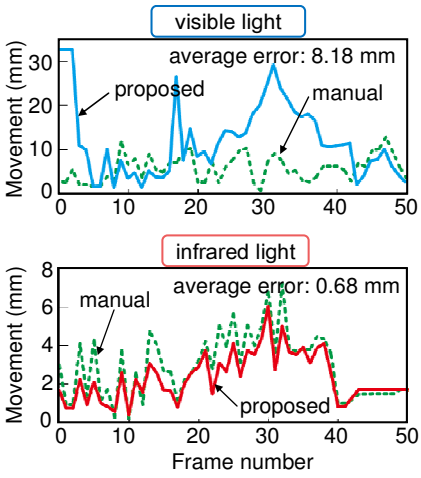

Fig. 11 Tracking accuracy of our system with the visible light or infrared light. The target object was the nose. The system outputs are compared with the results of manual tracking.

V. Conclusion

We developed a real-time contactless motion detection system with the GPGPU, and evaluated the performance of the system under the same conditions as those in actual radiation therapy. The results of our experiments showed that our system could detect the motion of the targeted objects within a mean error of 0.7 mm under the illuminance change. The results suggested the feasibility of its clinical application.

References

[1] T. Rohlfing, J. Denzler, C. Gräßl, D.B. Russakoff, C.R. Maurer "Markerless real-time 3-D target region tracking by motion backprojection from projection images," *IEEE Trans. Med. Imag.*, vol.24, pp.1455-1468, Nov. 2005

[2] H. Shirato, S. Shimizu, T. Kunieda, K. Kitamura, M. van Herk, K. Kagei, et al, "Physical aspects of a real-time tumor-tracking system for gated radiotherapy," *Int. J. Radiat. Oncol. Biol. Phys.,* vol.48, No.4, pp.1187-1195, 2005

[3] T.R. Willoughby, P.A. Kupelian, J. Pouliot, K. Shinohara, M. Aubin, M. Roaxh, et al. "Target localization and real-time tracking using the Calypso 4D localization system in patients with localized prostate cancer," *Int. J. Radiat. Oncol. Biol. Phys.* vol.65, No.2, pp. 528-534, 2006

[4] M.J. Murphy, "Image-guided motion adaptation in radiotherapy," in *Proc 4th IEEE Int. Symp. Biomed. Imag.*, pp. 996-999, 2007

[5] K. Ogawa, G. Mamiya, H. Iyatomi, Y. Oku, E. Kunieda, "Motion detection system with three USB cameras and active search algorithm for stereotactic radiosurgery", *in Proc. of World congress on medical physics and biomedical engineering (WC2009)*, vol.25, No.1, pp.37-40, Sep. 2009

[6] J. Nickolls, I. Buck, M. Garland, K. Skadron, "Scalable parallel programming with CUDA, " Queue, vol. 6, No. 2, March/April, 2008