# Semi-automated measurement of blood vessel diameter for arteriosclerosis retinae classification

Hirokazu Tachiki, Yuji Hatanaka, Susumu Okumura, and Kazunori Ogohara
School of Engineering the University of Shiga Prefecture
Hassaka-cho 2500, Hikone-shi, Shiga 522-8533, Japan

Ryo Kawasaki
Dept. Public Health, Yamagata University Faculty of Medicine
Iida-Nishi 2-2-2, Yamagata-Shi, Yamagata 990-9585, Japan

Koko Saito
Dept. Ophthalmology, Shinoda General Hospital
Sakura-cho 2-68, Yamagakata-Shi, Yamagata 990-0045, Japan

Chisako Muramatsu and Hiroshi Fujita
School of Medicine Gifu University
Yanagido 1-1, Gifu-shi, Gifu 501-1194, Japan

***Abstract*— The purpose of this study is to develop the measurement method of vein diameter near arteriovenous (AV) crossing for helping the diagnosis of AV nicking, which is one of the findings of arteriosclerosis retinae. We proposed a vein diameter measurement near AV crossing. However, our previous method failed to measure some diameters of veins, because the method mis-detected a vein as a tortuosity of one. In this study, blood vessel skeletons were smoothed by spline interpolation and P-type Fourier descriptor. Finally, we measured vein diameter by using the edge of the binarized blood vessel region obtained by our previous method. When the proposed method was applied to 20 retinal images, the average error and standard deviations was 1.7±1.4 pixel.**

***Keywords*— *arteriosclerosis retinae, AV nicking, measurement of diameter, spline interpolation, P type Fourier descriptor***

## I. INTRODUCTION

Arteriovenous (AV) nicking is significant findings for arteriosclerosis retinae. AV nicking is the phenomenon with compressed vein by a stiff artery on AV crossing. Therefore, changing the vein diameter near AV crossing is significant finding for AV nicking. Nguyen et al. have proposed a method for high accuracy classification of AV nicking [1]. However, they have defined a different index from the one that doctor use for diagnosis. We developed the vein diameter measurement near AV crossing to provide the physicians the measured vein diameter [2]. However, our previous method failed to measure some diameters of veins, because the method mis-detected a vein with a smooth curve as a tortuosity of one. Therefore, the purpose of this study is improvement of our previous method.

## II. METHODS

First, we input color retinal fundus image and manually specify the AV crossing point on this image. Next, we extracted blood vessel regions using black top-hat transformation combined with double ring filter [3]. The blood vessel skeletons were also extracted from the blood vessel regions, and they were divided into AV crossing and vein branches. The blood vessels were classified to artery and vein by using green channel pixels. The running direction of the blood vessel was extracted by the blood vessel skeleton. The blood vessel skeleton may have been mis-detected as tortuosity, so we smoothed the blood vessel skeleton by spline interpolation and P-type Fourier descriptor. Finally, we measured the length between the both edges of the vein region on the perpendicular of the vein-running-direction (P-LINE) as that diameter. We obtained the vein edge using two kind methods. The edge of the binarized vein regions was determined as a first method. As a second method, the vessel edge was detected by using zero-crossing method.

## III. RESULTS AND CONCLUSION

As evaluation, we calculated the average errors and standard deviations of two methods in 20 retinal images. Three researchers measured vein diameter manually, and the average of them was determined as the Gold Standard. The average error and standard deviation were 1.7±1.4 pixel and 2.6±3.1 pixel, respectively. The average of error by three researchers was 1.2 pixel, thus the proposed method may be useful although it should be improved slightly.

## REFERENCES

[1] U. T. V. Nguyen, A. Bhuiyan, L. A. F. Park, et al., "An Automated Method for Retinal Arteriovenous Nicking Quantification From Color Fundus Images, " IEEE Transactions on Biomedical Engineering, vol. 60, number.11, pp. 3194–3203, 2013.

[2] Y. Hatanaka, C. Muramatsu, T. Hara, et al., "Automatic arteriovenous crossing phenomenon detection on retinal fundus images," Proc. SPIE, Vol.7963, pp.79633V-1-8, 2011.

[3] C. Muramatsu, Y. Hatanaka, T. Iwase, et al., "Automated detection and classification of major retinal vessels for determination of diameter ratio of arteries and veins," Proc. SPIE, Vol.7624, pp. 76240J-1-8, 2010.

# Automated retinal blood vessel extraction using high-order local autocorrelation

Kazuki Samo,Yuji Hatanaka,
Susumu Okumura,and Kazunori Ogohara
School of Engineering, the University of Shiga Prefecture
Hassaka-cho 2500, Hikone-shi, Shiga 522-8533, Japan
oo23ksamo@ec.usp.ac.jp

Chisako Muramatsu and Hiroshi Fujita
Graduate School of Medicine, Gifu University
Yanagido 1-1, Gifu-shi, Gifu 501-1194, Japan

***Abstract*—The aim of this study is to develop an automated retinal blood vessel extraction using high-order local autocorrelation (HLAC). It does not require a preset model, thus it appropriates for a blood vessel extraction. We increase the number of mask patterns to improve performance. Black-top-hat (BTH) is used as pre-extraction and calculated HLAC from its output. The blood vessels are classified using an artificial neural network (ANN) with 105 HLAC features and 4 pixel-based features. The blood vessels are then classified using second ANN, which is inputted ANN and 4 features. As a result of the proposed method, the area under the curve based on receiver operating characteristics is 0.960.**

***Keywords—blood vessel extraction, retinal image, high-order local autocorrelation, black-top-hat, artificial neural network***

## I. INTRODUCTION

Retinas are useful for early detection of lesions such as hypertensive retinopathy. Thus, many blood vessel extraction methods have been proposed [1, 2]. We proposed a method using black-top-hat transform (BTH) [2], but that performance was not enough. On the other hand, a method based on the relation of neighbor pixels has not been proposed as far as we know. High-order local autocorrelation (HLAC) [3] is a feature extraction method depending on the autocorrelation between neighbor pixels. HLAC features is appropriate for center-shifted hotspot pattern feature extraction. Patten of blood vessels are so varied that designing a flexible model is not practical. Unlike a template-matching, HLAC does not require a preset model. Thus, HLAC is expected to be effective for blood vessel extraction.

## II. METHODS

The flow of the proposed method is as follows: The blood vessels are first enhanced by using BTH. Then, the region of interest (ROI) which is the area around the each pixels is defined, and four kinds of pixel-based-features and HLAC-based-features on ROI is calculated. Four features are maximum, minimum, average and standard deviation of the pixel values on ROI. A filter based on HLAC is defined using artificial neural network (ANN). All pixels are finally classified into a blood vessel and others by using a second ANN (ANN2) inputted 5 kinds of pixel values, which are HLAC filter (ANN), Gabor filter, double ring filter, BTH and green-channel component of color retinal image.

The $N$-th order HLAC is calculated by the following autocorrelation:

$$R_N^{\mathrm{T}}(a_1, a_2, \dots, a_N) = \sum_r I(r)I(r+a_1)\dots I(r+a_N) \qquad (1)$$

where $I(r)$ is a intensity of the pixel, $T$ is the local pattern number, $r = (x, y)'$ (the dash denotes the transpose) is position vector $a$, $a_N$ are the displacement vectors, and $x$ and $y$ are coordinates in the image. $N$ was limited to ($N$ = 0, 1, 2). A general HLAC has 35 mask patterns limited by reducing shift-invariant mask patterns when the range of displacements was also limited to within a local 3 × 3 window. In this study, we extend each of 35 mask's size from 3 × 3 to 5 × 5 and 7 × 7. Thus, we calculated total 105 HLAC features ($T$ = 1, 2 … 105). These HLAC features are weak to turned image, thus we use polar transform to ROI before calculating HLAC features.

## III. RESULTS AND CONCLUSION

In this study, we used images which is obtained from the "Digital Retinal Image for Vessel Extraction" (DRIVE) database, which includes 20 training images and 20 test ones [1]. In the output images of ANN, the blood vessel regions containing low contrast vessels have clearly white values. However, the blood vessel regions of the output images are thicker than that of the answer images. The outputs were evaluated using the area under the curve (AUC) based on receiver operating characteristics (ROC) analysis. As a result, the AUC of ANN was 0.935 and that of ANN2 was 0.960 as a result of our study. On the other hand, the AUC of our previous method using BTH is 0.918. Thus, our proposed method is effective for blood vessel extraction.

## REFERENCES

[1]. J.J. Staal, M.D. Abramoff, M. Niemeijer, M.A. Viergever, and B. van Ginneken, "Ridge based vessel segmentation in color images of the retina," IEEE Trans. Med. Img., vol.23, no.4, pp.501–509, Jun. 2004.

[2]. C. Muramatsu, Y. Hatanaka, T. Iwase, et al., "Automated selection of major arteries and veins for measurement of arteriolar-to-venular diameter ratio on retinal fundus images," Comput. Med. Imag. Graph., vol. 35, pp.472–480, 2011.

[3]. N. Otsu and T. Kurita, "A new scheme for practical, flexible and intelligent vision systems," In Proc. IAPR Workshop on computer Vision, pp.431–435, 1988.

# Classification of Diffuse Lung Diseases using Convolutional Neural Network for Practical Application

Noriaki Hashimoto, Shoji Kido and Yasushi Hirano
Graduate School of Medicine
Yamaguchi University
Ube, Japan
n_hashi@yamaguchi-u.ac.jp

***Abstract*— The diffuse lung diseases (DLDs) refer to a series of disorders that affect large pulmonary tissues in the lungs. Currently, several studies about classification of DLDs using CT images have been conducted, for example, where regions of interest in CT images are classified into normal tissues and six kinds of typical DLD patterns including consolidation, ground-glass opacity, honeycombing, reticular, emphysema and nodular.**

**Deep learning techniques are utilized in various fields, especially image processing. Obviously, they could be also applied to medical images for classification, object detection and so on. However, generally deep learning requires a huge number of training data for good result and shortage of training data would be often problem. To solve such problems, data augmentation in which original training data are artificially shifted, scaled or distorted to obtain the enough number of training images is often employed. Even with the above solution, there is a concern about computing time to learn increased dataset remaining.**

**In this study, we have implemented an algorithm for classifying DLD patterns of CT images with the ensemble learning using natural-image models to perform convolutional neural network with small dataset. As the result, we can get the good classification results compared to the conventional method by the proposed algorithm even with small CT data.**

***Keywords—convolutional neural network; diffuse lung disease; CT images; Classification***

# Precise extraction of liver tumor regions using both axial and coronal images of EOB-primovist MRI

Akihiko Satoh, Junichi Hasegawa
Graduate School of Computer and Cognitive Sciences,
Chukyo University
101 Tokodachi, Kaizu-cho, Toyota, Aichi, 470-0393 Japan
E-mail: h11405m@st.chukyo-u.ac.jp

Shigeru Nawano
Graduate School of Health and Welfare Sciences,
International University of Health and Welfare
1-4-3 Mita, Minato-ku, Tokyo, 108-8329 Japan

***Abstract*— This paper presents a procedure for precise extraction of liver tumor regions from hepatocyte phase images of EOB-primovist MRI (EOB-MRI). In the liver diagnosis, hepatocyte phase images can be used for presence diagnosis of liver tumors. For supporting this diagnosis by computer, we have already developed a procedure for extracting nodule regions as candidates of liver tumors from an axial image. In the paper, a new procedure to extract nodules more precisely by using both axial and coronal images is proposed. In this procedure, after nodules are extracted from each of axial and coronal images, both results are unified. In the experiments using 12 pairs of axial and coronal images, 25 pairs of corresponding nodule regions were extracted and each of them was unified into a single-shaped region. From the results, it was known that by using both axial and coronal images, nodule regions could be extracted more precisely in shape, compared with case of using either image.**

***Keywords*— CAD, liver, EOB-primovist MRI, hepatocyte phase image, axial and coronal images***

## I. INTRODUCTION

Great progress in medical imaging technology has made it possible to perform detailed diagnosis of the human body using 3D images such as CT and MR images. However, as the number of 3D images per patient was significantly increased, the burden of reading doctors became heavier. Therefore, computer-aided diagnosis (CAD) systems for 3D images have been needed for improving diagnostic accuracy and reducing reading burden [1-3].

Recently, MR images with EOB-primovist contrast agent are used as a new modality for liver diagnosis. EOB-primovist is a new fat-soluble contrast agent for MRI developed in 2008, and is easily incorporated into the hepatocytes, compared to water-soluble gadolinium contrast agent widely used so far. Particularly, in the hepatocyte phase at about 20 minutes after injection, EOB-primovist is taken into enough to normal liver cells, but not to other tissues (liver cell cancer, metastatic liver cancer, cysts, etc.) and blood vessels. Therefore, hepatocyte phase images can be effectively used for the presence diagnosis of liver tumors [4,5]. However, in case of reading tumors smaller than 5mm, supports by a CAD system have been still desired even for the hepatocytes phase images.

From this background, we have already developed a procedure for extracting nodule regions as candidates of liver tumors from axial images of the hepatocyte phase of EOB-primovist MRI [6]. This procedure consists of three steps; liver region extraction, vessel region extraction and nodule detection. However, shape preciseness and volume of detected nodules were insufficient due to low spatial resolution between slices of the input image.

In this paper, we propose a new procedure to extract nodules more precisely using both axial and coronal images of the hepatocyte phase. In this procedure, after nodule regions are extracted from each of axial and coronal images, corresponding regions are unified. By this, it is expected that each result can cover short of volume in the other.

In Section II, axial and coronal images are explained briefly, in Section III, the processing procedure is described, and in Section IV, experimental results are shown with some discussions.

## II. AXIAL AND CORONAL IMAGES

When taking *X*, *Y* and *Z*-axes to a human body as shown in Fig.1, a sectional image perpendicular to the *Z*-axis is called an axial image, and one perpendicular to the *Y*-axis called a coronal image.

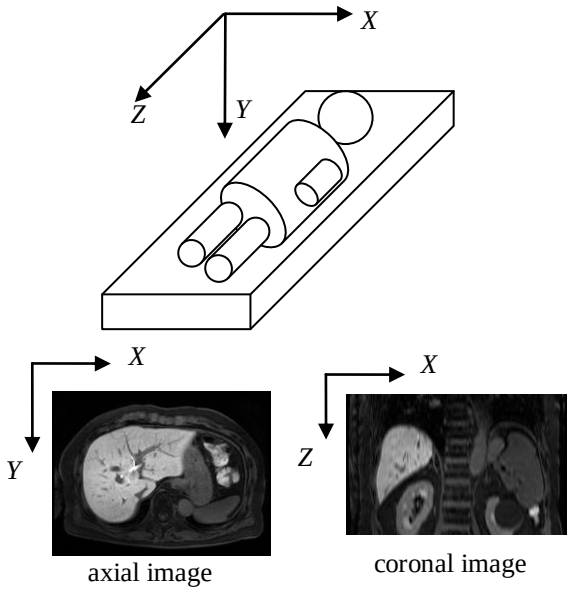



Fig. 1. Illustration of axial and coronal images

## III. PROCESSING PROCEDURE

The common procedure for extracting nodule regions from axial and coronal images is shown in Fig.2.

First, region growing followed by closing operation is applied to the input image for extraction of the liver region. The region growing process starts from a point in the liver region specified manually. The region is enlarged guradually by merging neighborhood pixels within a predetermined density range (thresholds T1, T2). This process is terminated when there is no pixel to be merged. Usually the result obtained above is a liver region with some missing areas (holes) caused by vessel regions or nodule regions. To fill up such holes, closing with the structure element of the radius R1 is performed.

Then, image subtraction between the liver region and the region growing result obtained above is performed. Usually the result includes both vessel regions and nodule regions. This is called "mixed regions".

Finally, to extract only nodule regions from mixed regions, opening with the structure element of the radius R2 is applied. The actual operation consists of shrinking, small and/or large component removal (volume thresholds T3, T4) and expansion.

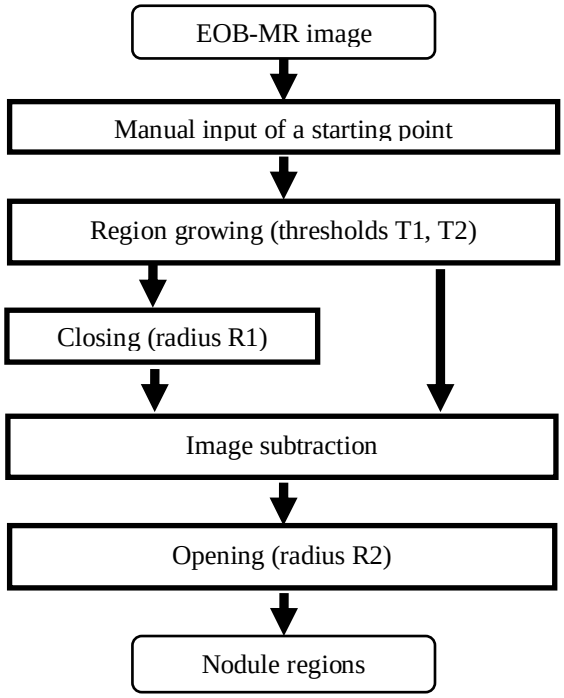



Fig. 2. Processing procedure

The procedure is applied to each of axial and coronal input images individually, provided that density values of the coronal image are transformed so as to have the same average value and the variance in the liver region as those of the axial image After nodule extraction from axial and coronal images individually, each pair of corresponding nodule regions is registrated with their centers of gravity, and then unified into a single-shaped region by taking logical addition of .two binary components

## IV. EXPERIMENTS AND DISCUSSION

### A. Experimental samples and parameter values

Sample images used in the experiment are 12 pairs of axial and coronal images of the hepatocyte phase of EOB-primovist MRI taken from 8 subjects in the International University of Health and Welfare Mita Hospital. Image sizes of axial and coronal images are 512 × 512 × 80 ~ 90 [voxel], 512 × 60 ~ 75 × 512 [voxel], respectively. The spatial resolution within a slice, the slice interval and the slice thickness are 0.694 ~ 0.949 [mm / voxel], 1.0 [mm], 2.5 [mm], respectively. Each of input images is resampled so that the spatial resolution is 1.0 [mm / voxel].

Density thresholds T1, T2, volume thresholds T3, T4 and radii R1, R2 of structure elements are 400, 700, 5, 30, 30 and 1, respectively. These values were determined experimentally.

### B. Results and discussion

First, nodule detection results from axial and coronal images individually are shown in Table 1. There were 42 nodules in 12 cases used in the experiment, and the average TP rate was 66.7% (=28/42) the same for both images. Suppose that if a nodule region can be detected from either the axial image or the coronal image, detection of the nodule is considered to be successfully done, the average TP rate becomes 73.8 % (=31/42) as shown in Table 2.

Table 1 Nodule detection results for each of axial and coronal images

| Image No. | Number of nodules | TP rate | |
|---|---|---|---|
| | | Axial | Coronal |
| 1 | 1 | 1/1 | 1/1 |
| 2 | 0 | 0/0 | 0/0 |
| 3 | 2 | 1/2 | 2/2 |
| 4 | 1 | 1/1 | 1/1 |
| 5 | 2 | 1/2 | 1/2 |
| 6 | 1 | 0/1 | 1/1 |
| 7 | 6 | 4/6 | 4/6 |
| 8 | 1 | 1/1 | 1/1 |
| 9 | 11 | 6/11 | 4/11 |
| 10 | 2 | 2/2 | 2/2 |
| 11 | 5 | 3/5 | 4/5 |
| 12 | 10 | 8/10 | 7/10 |
| Total | | 28/42 (66.7%) | 28/42 (66.7%) |

Table 2 Nodule detection results using both axial and coronal images

| Image No. | Number of nodules | TP rate |
|---|---|---|
| 1 | 1 | 1/1 |
| 2 | 0 | 0/0 |
| 3 | 2 | 2/2 |
| 4 | 1 | 1/1 |
| 5 | 2 | 1/2 |
| 6 | 1 | 1/1 |
| 7 | 6 | 4/6 |
| 8 | 1 | 1/1 |
| 9 | 11 | 6/11 |
| 10 | 2 | 2/2 |
| 11 | 5 | 4/5 |
| 12 | 10 | 8/10 |
| Total | | 31/42(73.8%) |

Next, the precise extraction procedure proposed here was applied to 25 nodule pairs were extracted successfully from both images. Results are shown in Fig.3, Fig.4 and Table 3.

Fig. 3 is an example of corresponding nodule regions extracted from axial and coronal images. In this case, both regions have almost the same volume, but their shapes are slightly different. Fig.4 is an example of unification of corresponding nodule regions. From 3D display in this figure, it is known that the shape of nodule regions obtained from each image is different, and that by unification, each region can cover short of volume in the other.

Table 3 shows examples of the shape and the volume of extracted nodule regions. From this, it can be seen that an elliptical (elongated) nodule has relatively large volume difference between regions extracted from each image, compared with spherical one, and that the volume of unified nodule region is larger than that of regions extracted from individual image. This suggests that the volume of nodules can be estimated more precisely by unification.

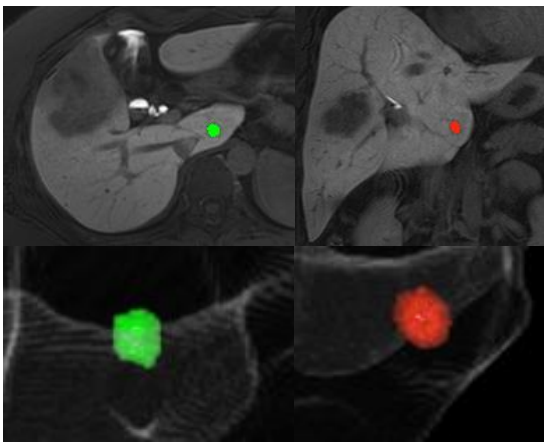


Fig. 4. An example of corresponding nodule regions extracted from each of an axial image (left) and a coronal image (right) and their 3D display

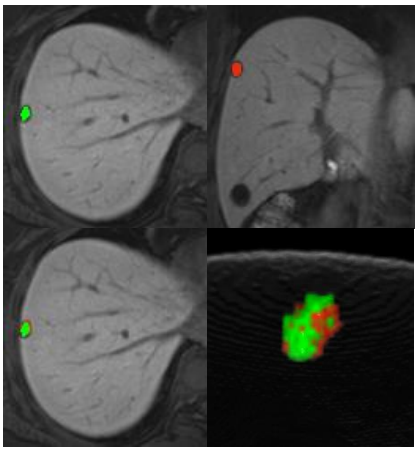


Fig. 3. An example of unifying corresponding nodule regions (upper left and right: nodules obtained from axial and coronal images, respectively, lower left and right: unified result and its 3D display, respectively

Table 3 Examples of shape and volume of extracted nodule regions

| Image No. | Nodule No. | Shape | Volume ($mm^3$) | | |
|---|---|---|---|---|---|
| | | | Axial | Coronal | Unified |
| 5 | 2 | spherical | 328 | 322 | 356 |
| 12 | 9 | spherical | 561 | 575 | 590 |
| 9 | 1 | elliptical (*Z*-direction) | 410 | 372 | 483 |
| 4 | - | elliptical (*Z*-direction) | 277 | 220 | 314 |
| 9 | 11 | elliptical (*Y*-direction) | 2287 | 2302 | 2639 |

## V. CONCLUSION

In the paper, a procedure for precise extraction of liver tumor regions using hepatocyte phase images of EOB-primovist MRI was proposed.

In the experiment, the performance of the previous procedure is evaluated first using 12 cases of actual MR images (axial and coronal images). The average TP rate was 73.8%.

Then, using 25 pairs of corresponding nodule regions extracted successfully from both images, the performance of the proposed procedure for precise extraction was evaluated. As a result, followings can be suggested.

(1) As the shape of the nodule is more elliptical (elongated), the volume difference between nodule regions extracted individually from axial and coronal images becomes larger.

(2) By unifying nodule regions extracted individually from axial and coronal images, the volume of nodules can be estimated more precisely.

Future work includes verification of above (1) and (2) with larger scale of samples, more effective use of both axial and coronal images for precise nodule extraction and development of a method for change detection of nodules from EOB-primovist MR images taken at different time from the same patient.

## ACKNOWLEDGMENT

Authors thank members of Hasegawa laboratory, Chukyo University for valuable discussion. In this work, the image media library MIST and the CAD platform PLUTO developed at Nagoya University were used. This work was supported in parts by the Grant-in-Aid for Scientific Research (B) from Japan Society for the Promotion of Science and by the Priority Research Project (P3) of the Knowledge Hub (Chi-no-Kyoten) from Aichi Prefecture, Japan.

## REFERENCES

[1] Yoshito Mekada: Progress in Computer Aided Diagnosis for Medical Images by Information Technology, ISSUE, 90(5): 406-411, 2007

[2] Masahiro Oda, Takayuki Kitasaka, Kensaku Mori, Yoshito Suenaga, Tetsyji Takayama, Hirotsugu Takabatake, Masaki Mori, Hiroshi Natari and Shigeru Nawano : Digital Bowel Cleansing Free Colonic Polyp Detection Method from Fecal Tagging CT Images, Trans IEICE, J91-D(7): 1094-1913, 2008

[3] Tomomi Takenaga, Yoshikazu Uchiyama, Toshinori Hirai, Hideo Nakamura, Yutaka Kai, Shigehiko Katsuragawa and Junji Shiraishi: Computer-Aided Detection of Metastatic Brain Tumors in Magnetic Resonance Images, JSRT, 69(6): 632-640, 2013

[4] Rumiko Kasai, Terumitu Hasebe, Nobuo Takada, Tsutomu Inaoka, Nobuyuki Hiruta and Hitoshi Terada: Comparison of Diagnostic Efficacy of Gd-EOB-DTPA-Enhanced MRI and Dynamic Contrast-Enhanced Multislice CT in Hepatocellular Carcinoma, Journal of Medical Society of Toho University, 59(6): 279-289, 2012

[5] Toshifumi Tada, Takashi Kumada, Hidenori Toyoda, K.Seiki, T.Kenji Takeshima, Sadanobu Ogawa, Katsuhiko Otobe, and N.Masayuki: Diagnosis of Hepatocellular Carcinoma by Contrast-Enhanced Ultrasound with Perfluorobutane and Enhanced Magnetic Resonance Imaging with Gd-EOB-DTPA, Kanzo, 51(3): 99-106, 2010

[6] Akihiko Satoh, Junichi Hasegawa, Shigeru Nawano: Development of a liver CAD system using hepatocyte phase images of EOB-primovist MRI, 2015 Joint Conference of IWAIT (International Workshop of Advanced Image Technology) and IFMIA (International Forum on Medical Imaging in Asia) (CD-ROM), No.357, 2015.01

# Ground Glass Nodule Detection by Chest Digital Tomosynthesis with Iterative Reconstruction Algorithm

Katsunori Miyata R.T.
Shiga University of Medical Science
Department of Radiology
Otsu, Shiga, Japan
katsu@belle.shiga-med.ac.jp

Satoru Matsuo R.T. PhD.
Kyoto College of Medical Science
Department of Radiological technology
Nantan, Kyoto, Japan

*Abstract*—To compare detectability of simulated ground glass nodules (GGND) on chest digital tomosynthesis (CDT) among 12 images obtained at 6 radiation exposure levels using 2 different reconstruction algorithms and analyze the influence of nodular size and its computed tomography attenuation value (CTAV) on GGND. CDT demonstrated sufficient GGND for less attenuated nodules with the diameter of 8mm or more even in the lowest radiation level (0.08mSv) and improved DS of GGN for more attenuated nodules with the diameter of 10mm at submilli-Sv with IRA. CDT has a sufficient potential to be used for detection of pure GGN and IRA synergistically boosts its ability.

*Keywords—component;* Digital Tomosynthesis, Iterative-Reconstruction algorithm(IR), Chest, GGND.

## I. PURPOSE

To compare detectability of simulated ground glass nodules (GGND) on chest digital tomosynthesis (CDT) among 12 images obtained at 6 radiation exposure levels using 2 different reconstruction algorithms and analyze the influence of nodular size and its computed tomography attenuation value (CTAV) on GGND

## II. MATERIALS AND METHODS

Seventy-four simulated GGNs (5, 8 and 10 mm in a diameter/ -630 and -800 Hounsfield of Unit (HU) in CTAV) were placed in the chest phantom with reproduced peripheral pulmonary structures in 14 different patterns of nodular distribution. In each of the 14 distribution patterns, 12 sets of reconstructed coronal images were obtained using CDT (SONIALVISION Safire, Shimadzu, Kyoto, Japan) with 6 different radiation level: 120kV-10mA/20mA/80mA/160mA, 100kV-80mA and 80kV-320mA (effective dose : 0.08/0.16/0.65/1.30, 0.39 and 0.77mSv in standard body habitus, respectively) with and without iterative dose reduction algorithm (IRA). Ten radiologists independently recorded GGN presence and their locations by continuously-distributed rating in total 168 sets of images. Receiver-operating characteristic (ROC) analysis was used to compare GGND of the 12 images in total and detection sensitivities (DS) of GGN were compared among the 12 images in subgroups classified by their nodular diameters and CTAV.

## III. RESULTS

In total 168 sets of images, GGND at 120kV-80mA with IR was similar to that at 120kV-160mA with IR, as area under ROC curve was 0.79±0.03 and 0.80±0.03, respectively, and higher than the other 6 images obtained at 120kV. ($p<0.05$) DS of GGN with the diameter of 8mm and -630HU in CTAV was 73.5±6.0 % in images at 120kV-10mA without IRA and similar to those in the other 11 images. ($p=0.157$) DS of GGN with the diameter of 10mm and -800HU in CTAV was 56.3±11.9 % in images at both 120kV-80mA and 120kV-160mA with IRA and higher than the other 4 images obtained at 120kV without IRA. ($p<0.05$)

## IV. CONCLUSION

CDT demonstrated sufficient GGND for less attenuated nodules with the diameter of 8mm or more even in the lowest radiation level (0.08mSv) and improved DS of GGN for more attenuated nodules with the diameter of 10mm at submilli-Sv with IRA. CDT has a sufficient potential to be used for detection of pure GGN and IRA synergistically boosts its ability.

## ACKNOWLEDGMENT

The authors acknowledge the great assistance of Ayumi Nitta-Seko M.D. PhD., Yoko Murakami M.D., Keiko Tsuchiya M.D. PhD., Akitoshi Inoue M.D., Sayaka Misaki M.D. (Shiga University of Medical Science, Otsu, Japan) and Khishigdorj Erdenee M.D. (Health Science University of Mongolia, Ulan Bator, Mongolia) for participating the simulated nodules detection study. The authors acknowledge the great assistance of Mr Mikio Moriyama and Mr Akira Kasai (Shimadzu co) for their technical support and important suggestions.

# Panoramic dental imaging based on a tomosynthesis approach

Taewon Lee
Department of Nuclear & Quantum Engineering
Korea Advanced Institute of Science and Technology
Daejeon, Korea
Leeto2000@kaist.ac.kr

Seungryong Cho
Department of Nuclear & Quantum Engineering
Korea Advanced Institute of Science and Technology
Daejeon, Korea
scho@kaist.ac.kr

***Abstract*— In dental panoramic systems, the teeth and jaws should be located in the central plane of the image layer. If the patient's teeth or jaws away from the central plane, the resolution of panoramic images rapidly decrease and the quality of the image degrade. In addition, since the shape and size of the jaws vary among individuals, some patients may not fit into a given panoramic system. The purpose of the study is to make a panoramic image by the tomosynthesis method. And also, automatically extract the regions are in focus in the panoramic tomosynthesis image to create a single layer**

# A hybrid approach for scatter correction in cone-beam CT

Jonghwan Min
Department of Nuclear and Quantum Engineering,
KAIST,
291 Daehak-ro, Yuseong-gu, Daejeon 305-701, South Korea

Seungryong Cho
Department of Nuclear and Quantum Engineering,
KAIST,
291 Daehak-ro, Yuseong-gu, Daejeon 305-701, South Korea
scho@kaist.ac.kr

***Abstract*— a beam-blocker composed of multiple strips is a useful gadget for scatter correction and/or for dose reduction in cone-beam CT (CBCT). We have developed a rebinned backprojection-filteration (BPF) algorithm for reconstructing images from the partially blocked cone-beam data in conjunction with the scatter correction. However, the proposed partial beam scan would additionally suffer from the incomplete data particularly in the off-midplanes with large cone-angles. Therefore, we modified beam blocked shape and reconstruction algorithm to reduce large cone angle artifacts. Additionally, we propose a hybrid model to compute scatter estimation by using the advantages of both measured and deconvolution kernel method. We demonstrated that the proposed method accurately estimates the scatter and reduced cone-beam artifact of image reconstruction**

***Keywords—Medical Imaging, Cone-beam CT***

## I. INTRODUCTION (*HEADING 1*)

## II. EASE OF USE

## III. PREPARE YOUR PAPER BEFORE STYLING

## IV. USING THE TEMPLATE

## REFERENCES

[1] G. Eason, B. Noble, and I.N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions,” Phil. Trans. Roy. Soc. London, vol. A247, pp. 529-551, April 1955. (*references*)

[2] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68-73.

[3] I.S. Jacobs and C.P. Bean, “Fine particles, thin films and exchange anisotropy,” in Magnetism, vol. III, G.T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271-350.

[4] K. Elissa, “Title of paper if known,” unpublished.

[5] R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.

[6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740-741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[7] M. Young, The Technical Writer’s Handbook. Mill Valley, CA: University Science, 1989.

# *Automatic Classification of Skull Deformity using Shape Analysis in 3D Head CT Images*

Min Jin Lee, Helen Hong*
Department of Multimedia Engineering
Seoul Women's University
Seoul, Republic of Korea
minjin, hlhong@swu.ac.kr

Kyu Won Shim
Department of Pediatric Neurosurgery
Yonsei University College of medicine
Seoul, Republic of Korea
SHIMKYUWON@yuhs.ac

Yong Oock Kim
Department of Plastic & Reconstructive Surgery
Yonsei University College of medicine
Seoul, Republic of Korea
sgm625@yuhs.ac

***Abstract*—Craniosynostosis is the premature fusion of the one or more cranial sutures resulting in skull deformity. 3D CT is the most standard imaging for early diagnosis and surgical planning of craniosynostosis. Despite the advances of 3D CT, cranial indices remain highly dependent on clinician experience. Thus, we propose a quantification method of the skull deformity using 2D shape descriptors representing the skull morphology in 3D head CT and classify the skull according to deformity type. First, cranial reference plane (R-plane) is established using frontonasal suture, opisthion and two porions. Running parallel to the R-plane, the position of reference planes in sagittal (S-plane) and bicoronal (C-plane) are then defined by ventricular landmarks within the brain. Second, skulls in S- and C-planes are segmented by 2D region growing with threshold values of 200HU. Then 2D shape descriptors are extracted from segmented skulls. A cranial index (CI) is calculated as the ratio of the cranial width divided by the cranial length. A cranial radius (CR) index is calculated as the minimum or maximum Fourier coefficient of the centroid distance function. The cranial partial slope (CPS) index is calculated as the average value of summated the centroid distance difference (CDD) at the forehead. A cranial extreme spot (CES) and near cranial extreme spot (NCES) indices are determined by the distribution of extreme spot in cranial chord matrix, which is the pairwise normalized Euclidean distance between cranial boundary. Finally, skulls are automatically classified using SVM into three types, sagittal synostosis, bicoronal synostosis and normal skull. The CPS index is shown a high accuracy with 93.33% in sagittal synostosis whereas the CES and NCES indices are shown a high accuracy with 98.89% in bicoronal synostosis. Our method can predict the type of craniosynostosis and quantify the condition of skull deformity***

***Keywords—shape descriptor; classification; craniosynostosis; 3D head CT image***

ACKNOWLEDGMENT

This work was supported by a research grant from Seoul Women's University (2015).

* Corresponding author. Email. hlhong@swu.ac.kr, Phone. +82 2 970 5756

# Automated Segmentation of Human Body Region Using a Line Profile in CT Scout Images

Byung Uk Ahn, Seung Won Shin, Kwang Gi Kim*
Biomedical Engineering Branch
National Cancer Center
Goyang, Republic of Korea
byunguk0227, sswcom98, kimkg@ncc.re.kr

Chang Hyun Lee
Department of Radiology
Seoul National University Hospital
Seoul, Republic of Korea
sswcom98@gmail.com

***Abstract*—In this paper, Our purpose is to automatically segment the four regions of the human body using a line profile obtained through image processing in CT scout images. We proposed a creation method of a line profile for a segmentation of human body. First, we used a median filter and OSTU thresholding method for noise reduction and optimization of the body region. Also, we applied the Contrast Limited Adaptive Histogram Equalization (CLAHE) method to emphasize the contrast between bright and dark regions. Second, we constructed a line profile with the median value obtained from each row of the optimal body region. And we carried out the characteristic point detection in the line profile for segmentation of human body. We tested 14 cases in 9 anterior-posterior(A-P) images and 5 lateral images. In order to verify results, we compared with the images of regions divided by a medical specialist. Chest region was detected in 11 of 14 images and pelvis region was detected in 5 of 9 images. The proposed method can be useful for the automatic segmentation of human body in CT scout images.**

***Keywords—image segmentation, image processing, CT scout image.***

# Detection of Clustered Microcalcification using Digital Mammography in Women With Dense Breasts for Early Breast Cancer

Jieun Oh, Kwang Gi Kim*
Biomedical Engineering Branch
National Cancer Center
Goyang, Republic of Korea
jieun12, kimkg@ncc.re.kr

Soo Yeul Lee
Bio-Medical IT Convergence Research Department
Electronics and Telecommunications Research Institute
Daejeon, Republic of Korea
seq@etri.re.kr

Eun Young Chae, Hak Hee Kim
Department of Radiology
College of medicine, Ulsan University
Seoul, Republic of Korea
chaeey@hanmail.net, hhkim@amc.seoul.kr

***Abstract*— To evaluate the validity of an automatic detection algorithm for microcalcification clusters through digital mammography in dense breasts. We proposed an automatic detection method for microcalcification clusters through digital mammography in dense breasts. First, we extracted the breast region, and then removed high-intensity artifacts, such as labels or scanning artifacts. Second, we enhanced the contrast to emphasize small microcalcifications in dense breast regions. Because the candidates for microcalcifications are observed as small bright blobs, we detected them by using a combination of the Laplacian of Gaussian (LoG) method and the Foveal algorithm. The candidates for the clustered microcalcifications, and their false positives were then reduced by using knowledge-based rules. We tested 112 cases with 234 malignant microcalcification clusters. The proposed method achieved a sensitivity of 97.43% at 0.27 false positives per image, and worked well with mammograms of various densities. The average detection accuracy in dense breasts was approximately 0.96. The proposed algorithm can be useful for the detection of microcalcification clusters in dense breasts for the early diagnosis of breast cancer.**

***Keywords—Mammography; Microcalcification; LoG; Foveal; Automatic Detection***

ACKNOWLEDGMENT *(Heading 5)*

We would like to acknowledge the financial support from the R&D Convergence Program of NST (National Research Council of Science & Technology) of Republic of Korea (Grant CAP-13-3-KERI).

# Can We Model the Confounding Physiophysical Factors in Quantitative CT Emphysema Index?

ChangYong Heo
Interdisciplinary Program in Radiation Applied Life Science,
Seoul National University College of Medicine
Seoul, Rep. Korea
heochy@snu.ac.kr

Jong Hyo Kim
Graduate School of Convergence Science and Technology
Seoul National University
Suwon, Rep. Korea
kimjhyo@snu.ac.kr

***Abstract*— Quantitative emphysema index measured with computed tomography has a potential to detect and assess the progression of emphysema noninvasively, and yet is known to suffer significant variability due to physiological and physical (physiophysical) confounding factors. We investigated the impact of various confounding physiophysical factors on CT emphysema index and explored the modeling of those factors in an attempt to compensate in CT emphysema quantification. We collected 335 CT scans from normal subjects (non-smoker, age 30-60 yrs) from a lung cancer screening database of our institution. Scan parameters were 40mAs, 120kVp, 1.0mm thickness, B30f with Siemens Sensation 16. We divided them into 325 training and 10 test data set. We obtained 10 additional scans of mild emphysema patients with identical scan parameters. The lungs, airways, and pulmonary vessels were automatically segmented, and two Emphysema Indices (EIs), RA950 and Perc15, were extracted from the segmented lungs using a software tool (SNU ImagePrism Pulmo). Additionally extracted from CT scans were two physiological factors such as total lung volume (TLV) and mode of lung attenuation (MoA) to reflect the inspiration level; and four physical factors such as effective body diameter and area (EBD, EBA), water equivalent body diameter and area (WBD, WBA) to reflect the body size-induced CT noise. The association of each physiophysical factor with EIs were obtained by Pearson correlation coefficients in training dataset. Then, we created a composite model reflecting the confounding relations of physiophysical factors with EIs using a logarithmic transform and multivariate regression. Finally, we evaluated the ability of our model to determine the likelihood of emphysema after compensation of confounding factors using the Z-score test. The correlation coefficient of physiological factors were 0.65 for TLV and 0.94 for MoA respectively in RA950, and 0.66 for TLV and 0.98 for MoA respectively in Perc15. The correlation coefficient of physical factors were 0.12 for WBD, 0.13 for WBA, 0.29 for EBD, and 0.28 for EBA, respectively with RA950; 0.02 for WBD, 0.03 for WBA, 0.20 for EBD, and 0.20 for EBA, respectively with Perc15. Our composite model produced much higher correlation coefficient of 0.976 for RA950 and 0.993 for Perc15. In test data set of 10 normal and 10 mild emphysema cases which could not be distinguished initially, our model produced Z-scores of 4.03±2.13 for the mild emphysema group, and 0.32±1.57 for the normal test group, providing the accuracy of 0.8 in detecting mild emphysema. Our study identified a combination of physiological and physical factors causing the variability in CT EIs, and could successfully model the composite their relationship. Our composite model has a potential to compensate the confounding physiophysical factors in quantitative emphysema assessment in CT.**

***Keywords—emphysema; low-dose; computed tomography; quantitative; physiophysical factors; modeling***

# 2D-to-3D Conversion for Laparoscopic Video in Minimally Invasive Surgery

Shih-Wei Huang[1], Guo-Shiang Lin[2]
[1]Chang Bing Show Chwan Memorial Hospital, Taiwan
[2]Department of Computer Science and Information Engineering, Da-Yeh University, Taiwan

Chang-Hau Chen[3], Kai-Che Liu[4], Wen-Nung Lie[3]
[3]Department of Electrical Engineering, National Chung Cheng University, Taiwan
[4]Medical Image Research Department, Asian Institute of TeleSurgery/IRCAD-Taiwan

***Abstract*—In this paper, we propose a model-based 2D-to-3D conversion scheme for laparoscopic videos in Minimally Invasive Surgery (MIS). MIS differs from traditional endoscopic examination in view of the existence of surgical instruments. It has been proved by experiments that with a 3D stereoscopic display, surgeons are capable of performing the surgical operations in a faster and more accurate manner. The proposed scheme is based on depth estimation from structure-from-motion (two consecutive frames are considered as the left and right views), alignment of the pre-constructed 3D instrument model with the estimated depths, and combination of the projected depths from the instrument model with the measured background (organ) depths. Our (instrument) model-based strategy is beneficial to depth conversion accuracy for the instruments. This presents a better 3D display quality which makes a surgeon perceive the relative position between the instrument and the background organ easier, hence speeding up and improving the accuracy of the surgical operations. Our simulation results show that the proposed scheme is capable of generating good depths for the instruments in MIS.**

***Keywords—2D-to-3D conversion; laparoscope; MIS***

## I. INTRODUCTION

Minimally Invasive Surgery (MIS) has been widely used for medical therapy for its less surgical risk and fast recovery for patients. In MIS, instruments and endoscopy are put into human's body via holes on, e.g., abdomen. Though MIS has some advantages, traditional monocular laparoscopic systems provide surgeons only 2D color videos which differ from what surgeons can see in the body. According to surgeons' experience, they can do instrument operations more fast and accurately with 3D depth perception. This is the motivation of converting 2D laparoscopic video into 3D stereoscopic one to provide surgeons a realistic environment for MIS. The technique concerning this goal is commonly called 2D-to-3D video conversion or stereoscopic video conversion [1].

So far, there are three kinds of 2D-to-3D video conversion techniques [1][2][3][4]. The first one is to analyze image contents and then directly generate binocular images. For example, Okino *et al*. [2] proposed a method called MTD (Modified Time Difference), by which motion analysis is performed for each frame and then a time delay is determined according to the magnitude of the motion vector. Two binocular images are generated by selecting two frames with the obtained time delay. However, MTD is suitable only for image sequences with horizontal object motions. Since surgeons move laparoscope arbitrarily, this kind of 2D-to-3D conversion methods is not suitably used in MIS.

The second one is that a depth map is first estimated and then utilized to create binocular images based on DIBR (Depth-Image-Based Rendering [1]). For instance, the work in [1] was characterized of estimating depths from multiple cues of an MPEG-decoded video, such as motion parallax, atmospheric perspective, texture gradient, linear perspective, and relative height. To deal with MPEG videos efficiently, the "motion" cue is extracted directly from the MPEG bit stream and others from the decoded frames. In their practical implementation, an input MPEG video is converted into a 2D+depth sequence or a stereoscopic image sequence, which can then be fed into modern stereocopic displays (for naked eyes or with glasses). In fact, due to the characteristics of laparoscopic images, depth cues are difficult to be estimated well. This kind of 2D-to-3D conversion methods may obtain poor depth maps, especially that the wrong depth relationship between the instrument tip and the organ will influence the surgeon's 3D perception and then the operations.

Recently, Kumar *et al*. [4] developed a model-based 2D-to-3D conversion method. In [4], a shape-from-shading algorithm is first used to generate a 3D shape of the organs in one frame. Then a manually selected region of 3D shape in this frame is registered with the 3D CT phantom model by using Iterative Closest Point (ICP) [6]. The depth information of the 3D CT model is combined with the laparoscopic 2D image for stereoscopic image generation. However, the method in [4] is not suitable for MIS due to the lack of surgical instruments that are unavoidable in MIS. The above observations motivate us to develop an (instrument) model-based 2D-to-3D conversion method for MIS.

## II. PROPOSED CONVERSION SCHEME

Figure 1 shows our conversion algorithm. Without considering estimating depths from a single image, we try to achieve this from two consecutive images. It is assumed that the surgeon moves the endoscopic camera slightly so that two consecutive images can be considered as the left and right views in a binocular configuration. After image rectification [7]

and disparity estimation [5], we can obtain an initial disparity (depth) map for the whole frame. By segmenting the instrument in color domain, the corresponding instrument depths can be identified.

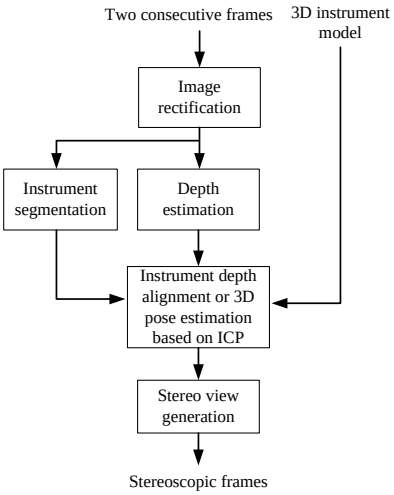



**Fig. 1** The block diagram of the proposed (instrument) model-based 2D-to-3D conversion scheme.

Due to the characteristic of laparoscopic images, e.g., low contrast, the estimated depths are likely to have errors, especially for the boundaries between instruments and organs. To have more accurate instrument depths for stereo display, a pre-constructed 3D instrument model is pose-aligned to the estimated instrument depths by using the ICP algorithm. The projected depths for the 3D instrument model are then used to replace the measured instrument depths accordingly. By depth replacement, the depth information for the instrument is refined to improve the stereo conversion quality.

### *A. Image rectification and disparity estimation*

In our work, we adopt a slightly moving camera to form a pair of images for disparity (depth) estimation. Though the camera motion is small, it might not be linear, causing the two consecutive frames not to be parallel. They should be rectified to improve the performance of disparity estimation. Here the laparoscopic frames are rectified based on an existing method [7]. An adaptive-window method [5] is used to estimate the disparity map, which is then post-refined with a maximum filter and a media filter to form $D_{stereo}$.

### *B. Instrument depth segmentation*

To register the 3D instrument model with the measured depths for 3D pose estimation, we first segment out the instrument from the depth image estimated above. Here we develop a multi-layer neural classifier for instrument segmentation from the color image. The hue (H) and the saturation (S) information for each pixel are used as the input features for classification. This relies on a pre-analysis that most surgical instruments are of gray or metallic colors which can be more easily distinguished from organs (often in maroon color) in the H and S domains. A three-layer feedforward neural network capable of non-linear feature mapping is implemented. The output neuron indicates the classification into instrument or non-instrument. Based on the trained neural classifier, pixels in the input frame can be classified and a mask map $\Omega^F$ indicating the instrument part can be obtained.

With the instrument mask $\Omega^F$, instrument depths $D_{inst}$ can be obtained accordingly.

### *C. 3D pose estimation for instrument*

We construct the 3D instrument model (composed of vertices and mashes, denoted as $M_{inst}$) via a 3D solid model software "metaseq240" (see Fig.2). As mentioned, the instrument depths can be estimated from every two consecutive laparoscopic images. It is now our goal to align them (i.e., estimating a translational and rotational matrix) so that the projection depths of the 3D instrument model can be used to replace the measured depths.

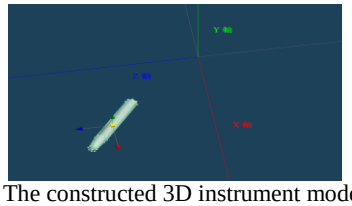


**Fig.2** The constructed 3D instrument model.

The position of the laparoscope is unknown at the beginning of MIS. This means that the transformation from 3D instrument model $M_{inst}$ to 3D instrument depths $D_{inst}$ is unknown. Assume that a virtual camera is placed at the origin of the 3D model space and the 3D instrument model $M_{inst}$ is placed somewhere. This transformation can be modeled as the product of a 3D rotation ($\boldsymbol{R}_{3D}$), a 3D translation ($\boldsymbol{t}_{3D}$), and perspective projection ($\boldsymbol{P}$). The projection matrix $\boldsymbol{P}$ is well-known to contain the intrinsic parameters which can be obtained after camera calibration. Here our goal is to estimate the right $\boldsymbol{R}_{3D}$ and $\boldsymbol{t}_{3D}$, which are capable of aligning $M_{inst}$ to $D_{inst}$ (or, making them most similar). The adopted technique to achieve this goal is the ICP algorithm.

Denoting each point in $M_{inst}$ as $[X_{3D}, Y_{3D}, Z_{3D}]^T$, its transformation to the depth domain can be described below:

(a) 3D geometrical transformation:

$$\begin{bmatrix} X'_{3D} \\ Y'_{3D} \\ Z'_{3D} \end{bmatrix} = \boldsymbol{R}_{3D} \cdot \begin{bmatrix} X_{3D} \\ Y_{3D} \\ Z_{3D} \end{bmatrix} + \boldsymbol{t}_{3D}, \tag{1}$$

(b) Projection:

$$\rho \begin{bmatrix} x_{2.5D} \\ y_{2.5D} \\ 1 \end{bmatrix}_{3\times1} = \boldsymbol{P}_{3\times4} \begin{bmatrix} X'_{3D} \\ Y'_{3D} \\ Z'_{3D} \\ 1 \end{bmatrix}_{4\times1}, \tag{2}$$

where $\rho$ is a scaling factor. Eq. (1)(2) can be used to obtain a projected depth image $D'_{inst}$ for the 3D instrument model at a certain viewing angle (determined by $\boldsymbol{R}_{3D}$ and $\boldsymbol{t}_{3D}$). We call

this projection result as 2.5*D* data since it is a depth image, but not true 3D data.

Now, we try to register the projected depth image $D'_{inst}$ to the measured depth image $D_{inst}$ by using ICP. The transformation between $D'_{inst}$ and $D_{inst}$ can be described as $\boldsymbol{R}_{2.5D}$ and $\boldsymbol{t}_{2.5D}$ in the 2.5*D* depth space:

$$\begin{bmatrix} x'_{2.5D} \\ y'_{2.5D} \\ z'_{2.5D} \end{bmatrix} = \boldsymbol{R}_{2.5D} \cdot \begin{bmatrix} x_{2.5D} \\ y_{2.5D} \\ z_{2.5D} \end{bmatrix} + \boldsymbol{t}_{2.5D}, \tag{3}$$

where $[x'_{2.5D} \quad y'_{2.5D} \quad z'_{2.5D}]^T$ denotes a point of $D'_{inst}$.

After configuring out $\boldsymbol{R}_{2.5D}$ and $\boldsymbol{t}_{2.5D}$ by ICP, we set $\boldsymbol{R}_{3D}$ = $\boldsymbol{R}_{2.5D}$ and then estimate $\boldsymbol{t}_{3D}$ accordingly. First, rotate the 3D instrument model (i.e., applying $\boldsymbol{R}_{3D}$ in Eq.(1)) via:

$$\begin{bmatrix} X''_{3D} \\ Y''_{3D} \\ Z''_{3D} \end{bmatrix} = \boldsymbol{R}_{3D} \cdot \begin{bmatrix} X_{3D} \\ Y_{3D} \\ Z_{3D} \end{bmatrix}. \tag{4}$$

then translate the result:

$$\rho \begin{bmatrix} x_{2.5D} \\ y_{2.5D} \\ 1 \end{bmatrix}_{3\times1} = \boldsymbol{P}_{3\times4} \begin{bmatrix} X''_{3D} + t_x \\ Y''_{3D} + t_y \\ Z''_{3D} + t_z \\ 1 \end{bmatrix}_{4\times1}, \tag{5}$$

where $\boldsymbol{t}_{3D} = [t_x \quad t_y \quad t_z]^T$. It is difficult to solve $\boldsymbol{t}_{3D}$ in Eq.(5) in an optimal sense. We vary the scaling factor $\rho$ by:

$$\rho_k = Z''_{3D} + 0.5 \cdot k, \; k = -6,-5,\cdots,0,\cdots 6, \tag{6}$$

where 0.5 is in the unit of centimeter (*cm*). For each given $\rho_k$, we can solve a unique $\boldsymbol{t}^i_{3D}$ for each 3D model vertex $v_i = [X^i_{3D}, Y^i_{3D}, Z^i_{3D}]^T$. These $\boldsymbol{t}^i_{3D}$'s can be averaged to obtain $\boldsymbol{t}_{3D,k}$ that corresponds to $\rho_k$. The final $\boldsymbol{t}_{3D}$ can be estimated by selecting the optimal $\rho_{k^*}$ that leads to the largest overlapping area between $D_{inst}$ and $D'_{inst}$. That is, we estimate the optimal $\boldsymbol{R}_{3D}$ and $\boldsymbol{t}_{3D}$ which will get a projected depth map $D'_{inst}$ mostly aligned to the real depth $D_{inst}$.

Note that the estimated $\boldsymbol{R}_{3D}$ and $\boldsymbol{t}_{3D}$ can be considered as the 3D pose parameters of the instrument model relative to the endoscopic camera, which might be useful in instrument pose tracking for Augmented Reality (AR) application of laparoscopic MIS. Since the initial pose of the 3D instrument model might be substantially different from the pose of the instrument in the captured image, the estimation of $\boldsymbol{R}_{3D}$ and $\boldsymbol{t}_{3D}$ should be made iterative. At iteration-1, we set $\boldsymbol{R}^n_{3D}$ to be $\boldsymbol{I}$ (Identity matrix) and $\boldsymbol{t}^n_{3D}$ to be $\boldsymbol{0}$ (zero vector), $n$=1, and compute the updated $\boldsymbol{R}^{n+1}_{3D}$ and $\boldsymbol{t}^{n+1}_{3D}$ according to $D'_{inst,n}$, where the superscript/subscript $n$ stands for the iteration number. That is, the pose parameters $\boldsymbol{R}^n_{3D}$ and $\boldsymbol{t}^n_{3D}$ at iteration $n$ are used in Eq.(1) for the computation of $\boldsymbol{R}^{n+1}_{3D}$ and $\boldsymbol{t}^{n+1}_{3D}$ from Eqs. (3)~(6). Once the condition that the overlapping between $D_{inst}$ and $D'_{inst,n}$ is above a threshold, the iteration can be stopped and the alignment between $M_{inst}$ and $D_{inst}$ is completed.

### *D. Foreground and background depth combination*

After convergence of finding $\boldsymbol{R}_{3D}$ and $\boldsymbol{t}_{3D}$, the projected instrument depth $D'_{inst}$ can be used to replace the real depth $D_{inst}$ for better 3D display quality. That is,

$$D(x,y) = \begin{cases} D'_{inst}(x,y) & (x,y) \in \Omega^F \\ D_{stereo}(x,y) & (x,y) \notin \Omega^F \end{cases}, \tag{7}$$

where $D(x,y)$ denotes the final depth value at $(x,y)$.

### *E. Stereoscopic frame generation*

According to $D(x,y)$ and the laparoscopic color frame, the left and right views for 3D stereoscopic display can be generated based on the well-known DIBR technique [1].

## III. EXPERIMENTAL RESULTS

To evaluate the performance of our proposed algorithms we select two clips of laparoscopic video sequences for testing. The frame size of each test image is 720×480 pixels. These test sequences are provided by IRCAD, Taiwan [8].

Figure 3 shows the results of the proposed scheme. The first row contains two input frames (#19 and #323). The second row gives the result of instrument segmentation (the non-instrument parts are set to be black). Experimentally, our neural classifier is capable of achieving 95% of classification accuracy. As shown from Fig.3 (c)(d), the instruments can be almost located except some boundaries. Some background areas might be mis-classified due to the existence of bright spots on the organ surfaces. Fortunately, their sizes are small and disconnected to the real instruments body and can thus be easily removed.

The third row of Fig. 3 shows the initial depth maps $D_{stereo}$ (brighter intensity means "near" distance) estimated based on the indicated frame and its preceding one. It is observed that the accuracy of the initial depth map is not quite good due to low contrast nature of the laparoscopic images. The parts with wrong depth estimations are marked with green circles, especially. Even, the depths within the interior of the instruments are not uniform (e.g., Fig. 3 (f)); the depths between the instrument/organ boundaries are not distinct (e.g., Fig. 3(e)). These defects make the quality of the initial depth maps unsuitable for 3D stereo display and 3D perception by the surgeons.

The fourth row of Fig. 3 gives the projected depths of the instrument model at the first iteration. It is clear that there is a substantial difference in the relative pose between Fig. 3 (e) and (g) and Fig. 3(f) and (h). The last row (i)(j) of Fig.3 shows the final depth maps after depth replacement. It is clear that after iterations, the projected depths and the measured depth can be aligned accurately, even with the poor instrument depths estimated at the beginning. Experimentally, our

algorithm is capable of converging in estimating $\boldsymbol{R}_{3D}$ and $\boldsymbol{t}_{3D}$ within 4~5 iterations (with an overlapping threshold of 80%).

Based on the final depth maps and the laparoscopic frames, Fig. 4 shows the synthesized left and right views in terms of interlaced format. They can be displayed on 3D display to perceive the 3D effect.

## IV. Conclusions and Future Work

In this work, a model-based 2D-to-3D conversion scheme for laparoscopic video is proposed for MIS application. Our scheme is featured of aligning a pre-constructed 3D instrument model to the measured depth image estimated via structure-from-motion technique, hence enabling the replacement of poor measured depths with the projected depths from 3D instrument model. Our scheme is capable of improving the depths of instrument itself and the depth edge between instrument and organ, thus presenting a more accurate depth perception of the instrument when the surgeons observe video on the 3D stereoscopic display.

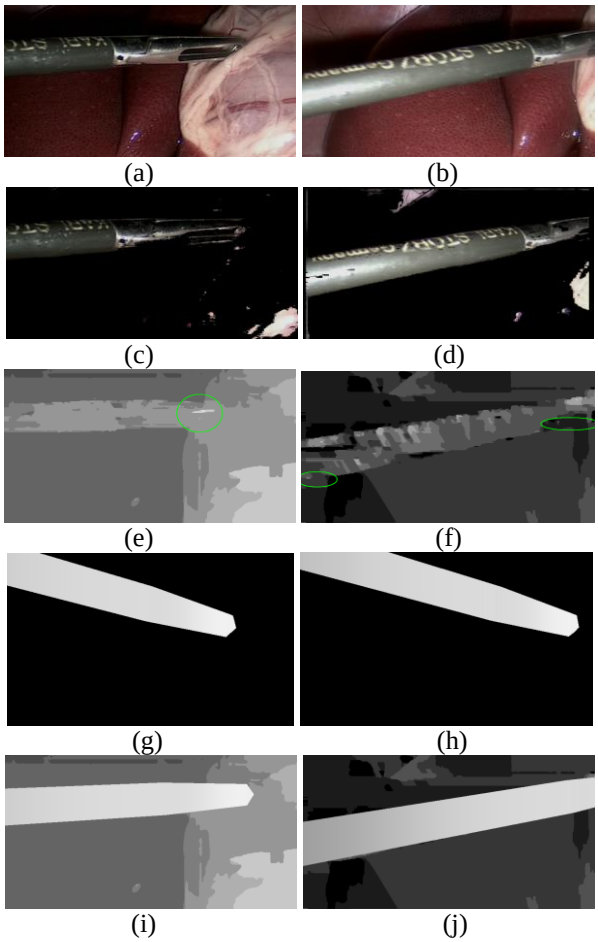



**Fig. 3** Experiments for two test sequences, (a)(b) input frame, (c)(d) instrument segmentation, (e)(f) initial depth maps, (g)(h) initial projected depths from 3D model, and (i)(j) final depth maps.

Our scheme is capable of not only improving the instrument depths for 3D perception, but also estimating the 3D pose parameters of the instrument relative to the endoscopic camera, which might be applicable to other applications such as Augmented Reality for laparoscopic MIS.

Compared with other prior works such as [4], our proposed method is featured of estimating the instrument depths more accurately and hence applicable to MIS, in addition to traditional endoscopic examinations.

Our future work would be to improve the quality of the initial disparity map. Moreover, we will evaluate the performance of the proposed method when two or more instruments exist in MIS.

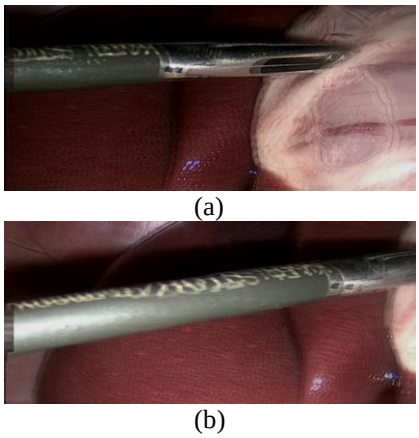



**Fig. 4** The synthesized left and right views in terms of interlaced format.

## References

[1] Guo-Shiang Lin, Hsiang-Yun Huang, Wei-Chih Chen, Cheng-Ying Yeh, and Wen-Nung Lie, "A stereoscopic video conversion scheme based on spatio-temporal analysis of MPEG videos," *EURASIP Journal on Advances in Signal Processing*, 2012.

[2] T. Okino, H. Murata, K. Taima, T. Iinuma, K. Oketani, "New television with 2D/3D image conversion technologies," *Proc. SPIE* 2653, 96–103, 1996.

[3] J. Zhou, Q. Zhang, B. Li, and A. Das, "Synthesis of stereoscopic views from monocular endoscopic videos," *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp. 55-62, 2010.

[4] A. Kumar, Y. Y. Wang, C. J. Wu, K. C. Liu, and H. S. Wu, "Stereoscopic visualization of laparoscope image using depth information from 3D model," *Computer Methods and Programs in Biomedicine*, Vol. 113, pp. 862–868, 2014.

[5] Guo-Shiang Lin, Chun-Yu Chen, Chun-Ting Kuo, and Wen-Nung Lie, "A computing framework of adaptive support-window multi-lateral filter for image and depth processing," *IEEE Trans. on Broadcasting*, Vol. 60, No. 3, pp.452– pp.463, 2014.

[6] P. J. Besl, N. D. McKay, "A Method for Registration of 3-D Shapes," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 14**,** No. 2, pp. 239–256, 1992.

[7] H. H. Wu, C. Y. Hung, P. K. Weng, and Y. Y. Wu, "Rectification of Stereoscopic Video Based on Affine Epipolar Geometry," *Proc. IPPR Conf. on Computer Vision, Graphics and Image Processing*, pp. 1608-1614, 2005.

[8] http://aits.tw/

# *Automatic Registration of Phalange Regions in CR Images Based on Deformation of the Likelihood*

*Shota Kajihara, Seiichi Murakami, Hyoungseop Kim, Joo Kooi Tan, Seiji Ishikawa*
Dept. Mechanical and Control Engineering
Kyushu Institute of Technology
1-1, Sensui, Tobata, Kitakyushu 804-8550, Japan
o344211s@mail.kyutech.jp

***Abstract*—** *Early detection and early treatment of the rheumatoid arthritis (RA) and osteoporosis are important task. However, some problems such as mass screening on data sets, mis-diagnosis are still remained in visual screening. In order to solve these problems and reduce the burden to physicians, needs of an automatic diagnosis system capable of performing quantitative analysis is anticipated. In this paper, we carry out the development of a registration method of phalanges regions from CR images of the hand to detect temporal changes of phalanges regions. The proposed method is carried out registration using salient region features (SRF) in image. We applied our method on 84 pairs of CR temporal images of phalanges regions and satisfactory experimental results are obtained.*

***Keywords—Image Registration; Sailent Region Features; Rheumatoid Arthritis; Osteoporosis; Computer Aided Diagnosis;***

## I. INTRODUCTION

Rheumatoid arthritis (RA) and osteoporosis are known as the main disease of the bone. Early detection and early treatment are important for these diseases. However, some problems such as mass screening on data sets and mis-diagnosis are still remained in visual screening [1, 2]. In order to solve these problems and reduce the burden to physicians, needs of an automatic diagnosis system capable of performing quantitative analysis is anticipated. In this research, we carry out the development of a registration method of phalange regions from CR images of the hand to perform a quantitative evaluation of RA and osteoporosis. The proposed method is carried out registration of phalanges regions using salient region features (SRF) [2, 3] in image. In addition, we introduce the deriving method of the deformation amount considering the likelihood of between SRFs.

## II. METHODS

We proposed a registration method using scale-invariant SRF. Our proposed method consists of three main steps (See Fig.1). In the first step called SRFD (Salient Region Feature Detection) step, circle area of the radius $s$ and center coordinates $X$ are set in the image. Then, the entropy is calculated for the circle area that does not exist in the background region. The value of radius is updated several times. The radius value at which the entropy becomes the maximum is set as the best scale $S_X$ and the saliency value $A(S_X,X)$ at $S_X$ is calculated. After computing at each center coordinates, the top $N$ in the saliency value is selected. In the second step called RCPM (Region Component Matching) step, pair $P_i$ is determined by the likelihood measurement based on the entropy correlation coefficient (ECC) for the SRF in each of the previous and the current image. Then, the top $M$ in the likelihood is selected. In the third step called RCFM (Region Configural Matching) step, registration of the entire image can be achieved by multiple binding of $P_i$.In this paper, the deformation amount $T$ are the rotation angle $\theta$ and the translation amount $t_x$, $t_y$ in the $x$ and $y$-direction. The deformation amount $T$ can be derived from a single pair. The translation amount is estimated from the difference between the central coordinates of the SRF which constitute the pair. The rotation angle indicates the angle at which the ECC is maximized. And, the global deformation amount is derived from pair joint $P^j$. The $P^j$ represents that it is binding with the $k$-th pairs.

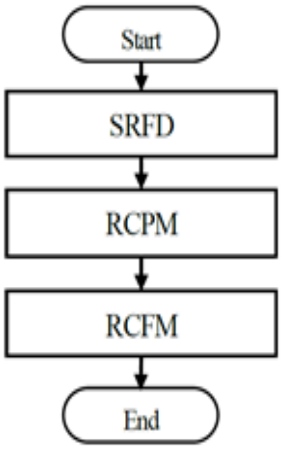



*Fig.1 Overview of the proposed processing*

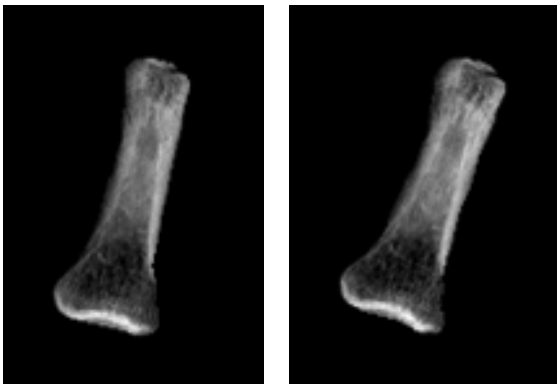
The global deformation amount $\hat{T}_k$ is derived by the equation (1).

$$\hat{T}_k = \frac{1}{L_{sum}} \sum_{i=1}^{k} L_i T_i \qquad (1)$$

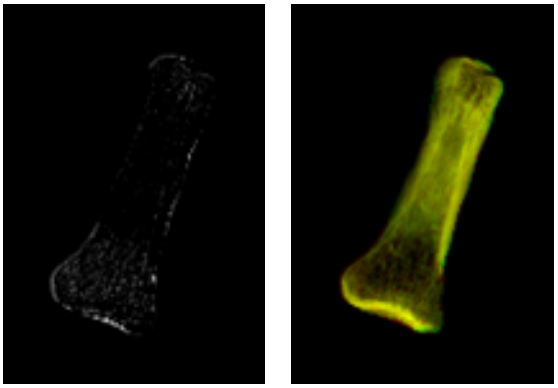
where $L_{sum}$ represents the sum of the likelihood of each pair. $L_i$ and $T_i=(t_x,t_y,\theta)$ represents the likelihood and the deformation amount of the pair $P_i$, respectively. Moreover, the alignment with the center of gravity takes place before performing the global deformation. By the introduction of the initial alignment with the center of gravity, we obtained the advantage that can respond to global deformation.

## III. EXPERIMENTAL RESULT

We applied our method on three pairs of CR temporal images of phalanges regions, which are consisted as the previous images and the current images from the same subject. The results of registration are shown in Fig.2, and TABLE 1.



(a) Previous (b) Current



(c) Subtraction (d) Fusion

*Fig.2 Experimental resutls*

TABLE I. EXPERIMENTAL EESULT

| | ***Conventional Method*** | ***Proposed Method*** | ***Combination*** |
|---|---|---|---|
| ECC | 0.597 | 0.616 | 0.619 |
| Times[s] | α + β 4.102χ· | (1) 3.724 ( | 140.75 |

In Table 1, combination represents a technique for searching for optimum deformation parameters from all the combinations of the deformation parameters. Value close as possible to the correct answer is derived by using this method. In Figure 2, (a) and (b) represents the previous and current image which is obtained same subject, respectively. Also, (c) and (d) represents the difference image and fusion image between deformed previous image and current image respectively.

## IV. CONCLUSION

In this paper, we proposed a method to reduce a registration error and processing time by use of scale-invariant SRF. From the experimental results, satisfactory experimental results were achieved. Therefore, our proposed method is effective for automatic registration of phalanges regions. As future works, we implement the features analysis of ROI using the registration results in order to perform quantitative evaluation of the disease. By using the aligned phalanges region, we expect that it enables automatic determination of the position of the region of interest and accurate characterization.

## REFERENCES


[1] J.Yoo, “Bone Age Estimation Method Using Normalized Shape Model”, IEICE Technical Report, pp.169-172, 2006.

[2] Y.Song, “A Hybrid Method for Reliable Registration of Digitally Reconstructed Radiographs and kV X-ray Images for Image-Guided Radiation Therapy for Prostate Cancer”, SPIE Medical Imaging, Vol.6918, 2008.

[3] X.Huang et al, “Hybrid Image Registration based on Configural Matching of Scale-Invariant Salient Region Features”, IEEE CVPR WS, 2004.