Automatic Registration of Phalanges Regions in CR Images Based on Scale-Invariant Salient Region Features

Shota Kajihara^{a*}, Seiichi Murakami^{a,b}, Hyungseop Kim^a
Joo Kooi Tan^a, Seiji Ishikawa^a

^aKyushu Institute of Technology 1-1, Sensui, Tobata, Kitakyushu 804-8550, Japan

^bDepartment of Radiology, University of Occupational & Environmental Health
1-1, Iseigaoka, Yahatanishi, Kitakyushu 807-8555, Japan

*Corresponding Author: kimhs@cntl.kyutech.ac.jp

Abstract

As two common orthopedic diseases, there are rheumatoid arthritis (RA) and osteoporosis. And early detection and early treatment are important for these diseases. However, some problem 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 by alignment between previous image and current one. The proposed method is carried out registration of phalanges regions using salient region features (SRF) in image. Specifically, SRF are detected from phalanges regions. Then, rigid registration technique is performed using deformation amount based on SRF. We applied our method on 84 pairs of CR temporal images of phalanges regions.

Keywords: Image Registration, Salient Region Features, Rheumatoid Arthritis, Osteoporosis, Computer Aided Diagnosis.

1. Introduction

Rheumatoid arthritis (RA) and osteoporosis are two common diseases. Dysfunction caused by joint destruction occurs when the symptom progress over a long period of time. Therefore, early detection of these diseases are important, to drug therapy, surgical

treatment, rehabilitation, and basic treatment. Frequency of the disease called osteoporosis increases bone if bone is degraded with aging. Incidences caused by these orthopedic diseases have increased dramatically in the context of the aging society of Japan. Due to the decrease of trabecular bone, fracture becomes more easily to happen. When a people decreased his trabecula it is difficult to recover. Therefore, discovery of the bone loss as early as possible is very important. Diagnostic imaging such as X-ray is mainly used in the diagnosis of osteoporosis and moreover diagnostic of imaging is the most important in the visual screening for the RA⁽¹⁻⁵⁾. However, experience of the physician is frequently required for visual screening. Evaluation of the structure of trabecular bone and measurement of bone mineral density is mainly used in the diagnosis of osteoporosis. However, visual assessment of trabecular bone structure by x-ray image is important because it cannot be diagnosed with osteoporosis. Diagnostic imaging is important for both osteoporosis and RA, but current diagnostic imaging is performed in the subjective assessment by the physician. Therefore, diagnostic results are different because of the experience of the physician interpreting the images. In addition, the burden to the doctors on account of the handling of the huge number of images manually also forms a problem^(6,7). In this paper, by motivation for a quantitative evaluation of RA and osteoporosis, we develop an automatic method for registration of phalanges regions to detect temporal changes by alignment between previous image and current one. The advantage of proposed method is that efficient registration is performed by using salient region

features(SRF) in two images respectively to determine the optimal deformation. We performed our methods on 84 pairs of CR temporal images of phalanges regions.

2. Methods

2.1 Overview of the proposed method

We proposed a registration method using scale-invariant SRF^(7,8). As general flow of processing, the proposed method consists of three main steps. At the first step, the salient region features is detected (SRFD) from the phalanges region which is automatic segmented from the CR images⁽⁹⁻¹¹⁾. As the second step, the region component matching (RCPM) is performed from the result of the SRFD step. In RCPM step, high correlation pairs are determined. And amount of deformation is determined for each pair. Finally, region configural matching (RCFM) is performed. In RCFM step, optimum amount of deformation is obtained by the binding of multiple pairs. The entire image is deformed by using the resulting amount of deformation (See Fig.1).

2.2 Salient Region Features Detection (SRFD)

We believe that every point regardless of their local characteristics (edgeness, cornerness, medialness, curvature, etc.) in the image can be made unique if a proper scale and its neighborhood is selected to calculate the feature. Therefore, we seek to use scale-invariant region features as the basis for our proposed registration method. In (12), a salient region feature detector is proposed. The salient regions are found using an entropy-based detector, which aims to select regions with highest local saliency. For each pixel x on an image, a probability density function (PDF) $p(s, x)$ is computed from the intensities in a circular region

of certain scale described by a radius s centered at x . The local differential entropy of the region is defined by:

$$H(s, x) = - \int_R p_i(s, x) \log_2 p_i(s, x) d_i \quad (1)$$

where i takes on values in the set of possible intensity values. The best scale S_x for the region centered at x is selected as the one that maximize the local entropy: $S_x = \text{argmax}_s H(s, x)$. Then the saliency value, $A(S_x, x)$, for the region with the best scale is defined by the extrema entropy value, weighted by the best scale and a differential self-similarity measure in the scale space:

$$A(S_x, x) = H(S_x, x) S_x \int_R \left\| \frac{\partial}{\partial s} p_i(s, x) |_{s_x} \right\| d_i \quad (2)$$

As the flow of SRFD step, firstly, for each pixel location x , the best scale S_x of the region centered and saliency value $A(S_x, x)$ are computed. Then, among the salient regions, the N most salient ones are picked as region features for the image. One of the main advantages of SRF is that they are theoretically invariant to rotation, translation.

2.3 Region Component Matching (RCPM)

Once we have extracted the salient region features from both the fixed and moving images, the alignment of the two images is achieved by finding a robust joint correspondence between multiple pairs of region features. This joint correspondence is then used to estimate the parameters of a desired transformation model. This translation can be described by three parameters : (t_x, t_y, θ) , where t_x, t_y are the translation along x and y directions respectively, and θ is the rotation angle. Also I_f is the fixed image, I_m is the moving image, and I_t is the transformed moving image. We aim to recover the parameters of a similarity transformation that geometrically transforms the moving image to be aligned with the fixed image. N_f and N_m are number of detected salient region features on I_f and I_m , respectively. $C_{i,j}$ is denoted the hypothesized correspondence between the i th region feature on I_f and the j th feature on I_m . And furthermore, $C_{i_1,j_1} \cap C_{i_2,j_2} \cap \dots \cap C_{i_k,j_k}$ is denoted that a hypothesized joint correspondence between multiple region feature pairs. In the RCPM step, we measure the likelihood of each hypothesized correspondence between a region feature from I_f and a region feature from I_m . That is to say, we want to measure the likelihood $L_{local}(C_{i,j})$ for each individual feature correspondence hypothesis $C_{i,j}$. We can then acquire a total ordering of these hypotheses according to their likelihoods. We define the likelihood to be proportional to the similarity between the interior intensities

Fig. 1. Overview of the proposed processing flow.

of the two salient regions involved. Let us denote the i th region on I_f as A and the j th region on I_m as B . The translation invariance is intrinsic by aligning the two region centers. To further achieve rotation invariance, we sample the parameter space for rotation sparsely and use the largest similarity value over all possible angles as the similarity between the two regions. The similarity measure we use is a normalized form mutual information, the Entropy Correlation Coefficient (ECC)^(13,14). The likelihood of a correspondence hypothesis $C_{i,j}$ is defined as:

$$\begin{aligned} L_{local}(C_{i,j}) &= \max_{\theta} (ECC(A, B^{\theta})) \\ &= \max_{\theta} \left(2 - \frac{2H(A, B^{\theta})}{H(A) + H(B^{\theta})} \right) \end{aligned} \quad (3)$$

where B^{θ} is the region B rotated angle θ , H indicates the joint or marginal differential entropy of the intensity value random variables of the two regions. We are able to sort these hypotheses in the order of descending likelihood. We then choose the top M such hypotheses to be used in the RCFM step to extract a globally consistent joint correspondence. The RCPM step also generates useful information regarding the transformation to align the two images, based on the purely local region-based matching. For instance, given a high likelihood correspondence between a region A on I_f and a region B on I_m , the rotation angle can be estimated as: $\theta = \arg\max_{\theta} ECC(A, B^{\theta})$. And the translation can also be estimated by the displacement between the center of the region A and the center of the region B . These estimates are associated with the related feature correspondence hypothesis, to provide the initial estimate for the transformation in the next RCFM step.

2.4 Region Configural Matching (RCFM)

In the RCFM step, we aim to detect a joint correspondence $C_{i_1,j_1} \cap C_{i_2,j_2} \cap \dots \cap C_{i_n,j_n} \dots$ between multiple pairs of region features, which results in the maximum likelihood in terms of global images “alignedness”. We measure the likelihood of a hypothesized joint correspondence with n feature pairs using the ECC measure between the overlapping portions of the fixed image I_f and the transformed moving image $T_m(I_m)$. Here the transformation $I_t = T_n$ is estimated from all feature pairs contained by current hypothesis. This can be written as:

$$\begin{aligned} L_{global} &= (C_{i_1,j_1} \cap C_{i_2,j_2} \cap \dots \cap C_{i_n,j_n}) \\ &= ECC(T_n(I_m), I_f) \\ (i_k, j_k) &\in [1, N_f] \times [1, N_m], k = 1, \dots, n \end{aligned} \quad (4)$$

In the end, we want to find a joint correspondence that has the maximum likelihood, while containing adequate number of feature pairs to recover the parameters of a similarity transformation. To address the combinatorial complexity in detecting the joint correspondence, we first compute a minimal correspondence and get an initial estimation of transformation. To choose this first correspondence, we measure $L_{global}(C_{i,j})$ for each individual feature match among the top M hypothesized correspondences resulted from the RCPM step. Using

Equation 4, the parameters of T_l are $(t_x, t_y, \theta)^{C_{i,j}}$ when measuring the likelihood of $C_{i,j}$. Then the first feature pair in the minimal correspondence base is the correspondence yielding the maximum likelihood, i.e.,

$$C_{i_1,j_1} = \arg\max_{C_{i,j}} L_{global}(C_{i,j}) \quad (5)$$

To allow converging to a globally optimal solution, we further use a generalized Expectation-Maximization (EM) algorithm to incrementally add in new feature pairs to the joint correspondence base, while refining the center locations of the corresponding features. Let current joint correspondence be $C = (C_{i_1,j_1} \cap \dots \cap C_{i_l,j_l})$. For each feature pair $C_{i,j}$ that is in the top M individual matches, but not in the current joint correspondence C , we estimate the likelihood of this feature pair being a valid correspondence in terms of global consistency as $L_{global}(C \cap C_{i,j})$, $C_{i,j} \notin C$. Then, we choose the new feature correspondence $C_{i,j}$ that has the maximum likelihood. We also require the addition of $C_{i,j}$ increasing the global image “alignedness”. If $L_{global}(C \cap C_{i,j})$ is larger than L_{global} , let the new joint correspondence be $C = (C \cap C_{i,j})$, and we re-compute the transformation T using the new joint correspondence. And furthermore we repeat EM step. If $L_{global}(C \cap C_{i,j})$ is smaller than L_{global} , we output current transformation T as the converged transformation to align the fixed image I_f and the moving image I_m . Having at most M iterations, our RCFM step is very efficient. Two key points contribute to this efficiency: First, pick a minimal correspondence base with only one feature pair; Second, use the EM algorithm to add in new feature pairs incrementally, thus enabling the converged joint correspondence to include as many good feature pairs as possible, while keeping a minimal complexity.

3. 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, Table 1. We had compared with the registration method using genetic algorithm (GA)⁽¹⁵⁾ and registration method using all combinations⁽¹⁵⁾. GA is the learning algorithm that engineered as to mimic process of evolution of creature. The registration method using GA is method that derive automatically optimum deformation. The registration method using all combinations is method that try all combination of deformation parameter. Value close as possible to the correct answer is derived by using this method. In the experimental results, the left and right hand side image represents the difference image and fusion image respectively. And Fig.3 shows the accuracy of registration. This graph shows the case where the expected value is maximized by the bond of six pairs. We evaluated the performance using the *ECC*(equation (3)), true-positive and false-positive. The descriptions of these two criteria are as follows.

$$TP = \frac{n(A \cap B)}{n(A)} \times 100[\%]$$

$$FP = \frac{n(C)}{n(B)} \times 100[\%]$$
(5)

Where A, B, C and n represent the target phalange region, registered region by our system, leak region, and number of

Fig. 2. Registration result.

Table 1. Performance evaluation.

	Using GA	Proposed Method	Combinations
ECC	0.604	0.615	0.621
Times	56.24	4.016	136.17

Fig. 3. Accuracy of TP and FP.

pixels respectively. *TP* represents the degree of overlap of the extraction region for the target phalange region. *FP* represents the percentage of leak region for the registration result. In this study, good performance with *ECC* of 0.615 and *TP* of 99.2[%] were obtained. Using our proposed method, the computational time for registration is reduced without reducing the registration accuracy.

4. Conclusions

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, good performance was achieved. Therefore, our proposed method is effective for automatic registration of phalanges regions. Furthermore, the processing time of our method is less than the GA. We can reduce the processing time about 1/14 compare with the GA algorithm. We think that it is because only the appropriate information that required for the registration was used without search of the optimal solution by evaluating the entire image. However, there are some cases of end of registration due to falling into local minimum. In other words, it is that the *ECC* is not increased. As the cause, derivation of the amount of deformation is considered. To overcome this problem, we should introduce another method for deriving the amount of deformation such as weighting considering the likelihood.

Acknowledgment

This work is partly supported by the Kitakyushu Foundation for the Advancement of Industry Science and

Technology, which is hereby gratefully acknowledged. Also the authors would like to thank MD. Takatoshi Aoki from the University of Occupational and Environmental Health, Japan for his constructive feedback and kind supports.

References

- (1) C.Alves et al, “Periarticular osteoporosis : a useful feature in the diagnosis of early rheumatoid arthritis? Reliability and validity in a cross-sectional diagnostic study using dual-energy X-ray absorptiometry”, *Rheumatology*, Vol.50, No.12, pp.2257-2263, 2011
- (2) N.Magarelli et al, “MR imaging of atlantoaxial joint in early rheumatoid arthritis”, *Radiol Med*, Vol.115, No.7, pp.1111-1120, 2010
- (3) M.Ogasawara, “Usefulness of Joint Ultrasonography in Medical Care for Rheumatoid Arthritis”, *Juntendo Medical Journal*, Vol.60, No.2, pp.107-111, 2014
- (4) D.Kyburz et al, “The long-term impact of early treatment of rheumatoid arthritis on radiographic progression : a population-based cohort study”, *Rheumatology*, Vol.50, No.6, pp.1106-1110, 2011
- (5) D.Topfer, “Segmentation and quantification of bone erosions in high-resolution peripheral quantitative computed tomography datasets of the metacarpophalangeal joints of patients with rheumatoid arthritis”, *Rheumatology*, Vol.53, No.1, pp.65-71, 2014
- (6) J.Yoo, “Bone Age Estimation Method Using Normalized Shape Model”, *IEICE Technical Report*, pp.169-172, 2006
- (7) 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
- (8) X.Huang et al, “Hybrid Image Registration based on Configurational Matching of Scale-Invariant Salient Region Features”, *IEEE*, 2004
- (9) S.Kajihara et al, “Automatic Segmentation of Phalanges Regions on CR Images Based on MSGVF Snakes”, *ICCAS*, 2014
- (10) J.Tang et al, “Vessel Boundary Tracking for Intravital Microscope Via Multiscale Gradient Vector Flow Snakes”, *IEEE Transaction Biomed. Eng.*, Vol.51, No.2, pp.316-324, 2004
- (11) M.Unser et al, “B-Spline Signal Processing : Part1-theory”, *IEEE Transaction on Signal Processing*, Vol.41, No.2, pp.821-832, 1993
- (12) T.Kadir and M.Brady, “Saliency, scale and image description”, *International Journal of Computer Vision*, Vol.45, No.2, pp.83-105, 2001
- (13) F.Maes et al, “Multi-modality image registration by maximization of mutual information”, *IEEE Transactions on Medical Imaging*, Vol.16, No.2, pp.187-198, 1997
- (14) H.Luan et al, “Multi-modal Image Registration by Quantitative-Qualitative Measure of Mutual Information (Q-MI)”, *CVBIA, LNCS 16*, pp.378-387, 2005
- (15) S.Kajihara et al, “Automatic Registration of Phalange Regions in CR Images Based on Genetic Algorithm”, *BMIRC*, 2015.