Quality Assessment Based on PCJO for Low-dose CT Images

Tianjie Xu, Lu Zhang, Yang Chen*, *Member*, *IEEE*, Huazhong Shu, *Senior Member, IEEE*, Limin Luo, *Senior Member*, *IEEE*

Abstract: The anthropomorphic model observer (MO) has played an important role in the assessment of medical imaging systems due to its tasked-based features. In this paper, an improved clinical-relevant anthropomorphic MO is proposed for the task of detecting and locating lesions based on low-dose computerized tomography (LDCT) and the perceptually relevant channelized joint observer (PCJO). During our former study, PCJO has been proposed for detecting and locating multiple hyposignals and hypersignals of conventional dose with unknown amplitude, orientation, size and location. In this paper, we extend PCJO to LDCT field to explore its generality. Experiments were conducted using different image sets obtained by two LDCT image reconstruction algorithms: filtered back projection (FBP) and adaptive statistical iterative reconstruction (ASIR), and seek to compare their advantages and disadvantages. Preliminary results show that the extended PCJO can detect and locate multiple lesions with unknown amplitude, orientation, size and location in LDCT reconstruction images and approach experts' performance under certain VDP threshold and combinations of channels.

Index Terms—PCJO, LDCT, image reconstruction.

I. INTRODUCTION

LDCT can significantly lower the risk of serious illnesses [1], but at the cost of damaging image quality because of increased noise and artifacts. Therefore, effective methods are required in order to enhance LDCT image quality [2], and better algorithms of image quality assessment should be proposed for the purpose of evaluating quality more precisely.

Although subjective assessment is reliable, there exists problems of long working time and prone to be influenced by subjective and objective factors. Thus accurate objective evaluation algorithms are necessary to conduct auxiliary assessment so as to reduce the burden of doctors.

Many modern objective image quality assessment algorithms are based on human visual system (HVS) that considers the visual system of human the best visual information receiving and processing system [3]. Therefore they include the sensation stage of the human vision processing and many psychophysical characteristics such as sub-band decomposition and contrast sensitivity function (CSF).

Another category of objective image quality assessment algorithms are based on model observers (MO). MOs mainly abstract a statistical notion from the image quality, represent characteristics of the image，background and noise in the form of mathematical expressions and describe the statistical notion using statistical decision theory[4][5].

As a typical anthropomorphic MO derived from the channelized joint observer (CJO) [6], the perceptually relevant CJO (PCJO) [7] can detect and locate multiple hyposignals and hypersignals of conventional dose with unknown amplitude, orientation, size, location and numbers of signals on single-slice images. The most striking difference between CJO and PCJO is the perceptual stage of PCJO that absorbs the sub-band decomposition characteristic of HVS and the task-based paradigm which considers image quality should be assessed in the context of a specific diagnostic task [5]. In order to approach human performance, the PCJO involves an HVS model named visual difference predictor (VDP) [8], on the output of which we apply a detection probability threshold T_p.

PCJO contains two main parts. The first part is the CJO training phase that trains the channelized parametric signal using the joint detection and estimation theory and the channelized method. The second part is the testing phase that can be divided into two sections. The first section is candidate selection that performs the localizing task. The candidate selection generates several test blocks for each image under test, and outputs the center position of each testing block on the image. The second part is the application of the CJO testing on candidates that performs the detection task. It calculates a test statistic for each testing block in order to judge the existence of lesions and then estimate parameters of lesions if they exist [7].

In this paper, we make efforts to apply PCJO to LDCT field in order to validate its generality and improve its clinical relations. Meanwhile, image sets based on different image reconstruction algorithms (see Fig. 1) are used: FBP and ASiR.

The rest of this dissertation is organized as follow: Section II introduces the mathematical model of background and signal and define the signal detection model for hyposignals; Section III introduces the PCJO image quality assessment experiments in detail; Section IV introduces the experimental results and makes a discussion; Section V concludes the paper.

(a)   (b)
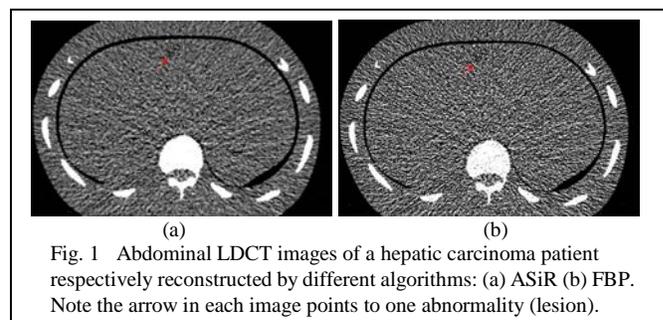
Fig. 1 Abdominal LDCT images of a hepatic carcinoma patient respectively reconstructed by different algorithms: (a) ASiR (b) FBP. Note the arrow in each image points to one abnormality (lesion).

II. MATHEMATICAL BACKGROUND

In this work, we use the same background and signal models as in the previous study [9]. In the task of detecting and locating hyposignal, the problem can be seen as the validation of one of the two following exclusive hypotheses:

$$H_k : \boldsymbol{g} = -k\boldsymbol{x}_\Lambda + \boldsymbol{b}, \quad k = 0,1 \tag{1}$$

where $\boldsymbol{g}$ is an $M \times 1$ column vector representing the digital image with M pixels, and the binary variable k controls the absence or presence of the signal x. $\boldsymbol{x}_\Lambda$ denotes a particular signal with signal parameters vector Λ.

Based on the joint detection and estimation theory, the signal parameter Λ and the hypothesis H_k are estimated to maximize the joint posterior probability $P(\Lambda, H_k / \boldsymbol{g})$:

$$\widehat{(\Lambda, H_k)} = \arg\max_{\Lambda, H_k} P(\Lambda, H_k / \boldsymbol{g}) \tag{2}$$

Using the same way as in [9] did, the estimation of Λ is:

$$\hat{\Lambda} = f(\Lambda_e)$$
$$= \begin{cases} \Lambda_e & -\boldsymbol{x}_\Lambda^t \boldsymbol{\Sigma}_b^{-1}(\boldsymbol{g} + \frac{1}{2}\boldsymbol{x}_\Lambda) > 0 \\ \text{any value in } \Re(\Lambda) & \text{otherwise} \end{cases} \tag{3}$$

Given the estimated parameter $\hat{\Lambda}$, we have:

$$P(H_k \mid \hat{\Lambda}, \boldsymbol{g}) \propto P(H_k) P(\boldsymbol{g} \mid H_k, \hat{\Lambda}) \tag{4}$$

The classical decision approach, based on signal detection theory, is to choose H_0 when $P(H_0 \mid \hat{\Lambda}, \boldsymbol{g}) > P(H_1 \mid \hat{\Lambda}, \boldsymbol{g})$, which means that the hyposignal is absent; otherwise we choose H_1, which means the hyposignal is present. So we have:

$$-\frac{1}{2}(\boldsymbol{g} + \boldsymbol{x}_{\hat{\Lambda}})^t \boldsymbol{\Sigma}_b^{-1}(\boldsymbol{g} + \boldsymbol{x}_{\hat{\Lambda}}) + \frac{1}{2}\boldsymbol{g}^t \boldsymbol{\Sigma}_b^{-1}\boldsymbol{g} \underset{H_0}{\overset{H_1}{\underset{<}{>}}} \ln\frac{P(H_0)}{P(H_1)} \tag{5}$$

$$\lambda = -\boldsymbol{x}_{\hat{\Lambda}}^t \boldsymbol{\Sigma}_b^{-1}(\boldsymbol{g} + \frac{1}{2}\boldsymbol{x}_{\hat{\Lambda}}) \underset{H_0}{\overset{H_1}{\underset{<}{>}}} \ln\frac{P(H_0)}{P(H_1)} \tag{6}$$

where λ is the test statistic for deciding the presence or absence of the hyposignal.

As in [10], we solve the dimensionality problem using the channelization method by introducing channel matrix. Then the estimation and the detection can be rewritten as:

$$\Lambda_e = \arg\max_\Lambda \frac{-1}{\|U(A_\Lambda')^t\|_F^2}(\boldsymbol{x}_0')^t (\boldsymbol{\Sigma}_b')^{-1}(A_\Lambda'\boldsymbol{g}' + \frac{1}{2}\boldsymbol{x}_0') \tag{7}$$

$$\hat{\Lambda} = f(\Lambda_e) \tag{8}$$

$$\lambda = \frac{-1}{\|U(A_{\hat{\Lambda}}')^t\|_F^2}(\boldsymbol{x}_0')^t (\boldsymbol{\Sigma}_b')^{-1}(A_{\hat{\Lambda}}'\boldsymbol{g}' + \frac{1}{2}\boldsymbol{x}_0') \tag{9}$$

where $\|U\|_F^2$ is a channel matrix energy normalization factor and $\|\cdot\|_F^2$ denotes the Frobenius norm of the matrix. $\boldsymbol{g}'$ is the channelized image: $\boldsymbol{g}' = U_\Lambda \boldsymbol{g}$. A_Λ' serves to map the channelized parametric signal $\boldsymbol{x}_\Lambda'$ to a channelized reference signal $\boldsymbol{x}_0'$: $A_\Lambda'\boldsymbol{x}_\Lambda' = \boldsymbol{x}_0'$. The matrices U and A_Λ' remain the same as in [7], where U is constructed by K steerable channels and J scale-shiftable channels. $\boldsymbol{\Sigma}_b'$ is the same as in [9].

III. EXPERIMENTAL PROTOCOL

A. Experimental images and parameters setting

To study the PCJO task performance, we selected the hepatic carcinoma as the studied pathology and LDCT as the studied modality. In order to compare the performances of different LDCT image reconstruction algorithm, we selected image sets reconstructed by the reconstruction algorithms FBP and ASiR respectively. From a database of the Nanjing First Hospital, we collected abdominal LDCT images of 10 healthy subjects and chose 200 independent non-contiguous images as the reference images. We then added subtractive lesions with an amplitude value between 33 and 35 to simulate the hepatic carcinoma lesions. The amplitude, orientation, size, positions and number of the lesions were random. For CJO training, 600 training blocks (300 without signal and 300 with signal) were extracted from healthy abdominal LDCT images, respectively. Other parameters were given in [9].

B. Subjective experiments

In order to get radiologists' performances, we conducted the same subjective experiment as in [7] with four radiologists including two CT experts (E1, E2) and two normal radiologists (R1, R2). Each radiologist performed the joint detection-localization task without time limit.

C. Performance evaluation

The Figure of Merit (FOM) was the area under the JAFROC1 curve (AUC) using the software RJafroc [11], which can show better performance in the detection and location task [12]. The significance level was set to 0.05 in this paper.

IV. RESULTS AND DISCUSSIONS

A. Results of radiologists & PCJO for FBP LDCT images

Table I

The task performance of all radiologists (TP means true positive mark and FP means false positive mark)

	E1	E2	R1	R2
JAFROC1 FOM	0.7055	0.7023	0.7022	0.7021
Standard Error	0.0230	0.0235	0.0132	0.0131
Detection rate	0.6226	0.6198	0.4105	0.4077
Number of TP marks	226	225	149	148
Number of FP marks	126	132	2	1
Total marks	352	357	151	149
Total lesions	363	363	363	363

For the detection-localization of simulated hepatic carcinoma lesions on abdominal FBP LDCT images, the JAFROC1 FOMs of four radiologists are calculated and shown in Table I. We observe that the FOMs of experts are similar to those of normal radiologists, while the detection rate of experts are significantly higher than those of normal radiologists.

Table II
Number of TP marks and FP marks under different threshold Tp

T_p	TP	FP
0.1	361	119
0.2	356	112
0.3	345	88
0.4	329	74
0.5	298	63
0.6	266	46
0.7	235	35
0.8	176	20
0.9	121	11

For the PCJO, we firstly study the influence of the probability threshold (T_p) by varying it from 0.1 to 0.9 with a step of 0.1. The numbers of TP and FP marks obtained by different T_p settings are shown in Table II. We can see that
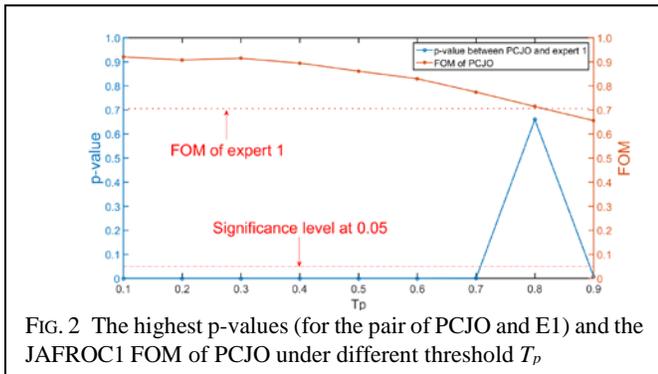
FIG. 2 The highest p-values (for the pair of PCJO and E1) and the JAFROC1 FOM of PCJO under different threshold T_p

when T_p is between 0.6 and 0.8, the TP and FP numbers of the PCJO seem close to those of radiologists.

We then further find the relation between the *p*-value (for examining the JAFROC1 FOM difference between PCJO and E1 with highest FOM and detection rate) and the probability threshold (T_p), as shown in Fig. 2. Note that for different thresholds, we tried different combinations of the number of steerable channel *K* and the number of scale-shiftable channel *J* to get the highest *p*-values. We see that when $T_p = 0.8$, there is no significant difference between the JAFROC1 FOM of PCJO and that of E1. In order to make the JAFROC1 FOM of PCJO closest to E1, we retain the value 0.8 for the T_p for the abdominal FBP LDCT images.

B. Results of radiologists & PCJO for ASiR LDCT images

Table III
The task performance of all radiologists (TP means true positive mark and FP means false positive mark)

	E1	E2	R1	R2
JAFROC1 FOM	0.7799	0.7781	0.7222	0.7151
Standard Error	0.0186	0.0184	0.0207	0.0217
Detection rate	0.6143	0.6584	0.6667	0.6198
Number of TP marks	223	239	242	225
Number of FP marks	71	61	64	88
Total marks	294	300	306	313
Total lesions	363	363	363	363

For the detection-localization of simulated hepatic carcinoma lesions on abdominal ASiR LDCT images, the JAFROC1 FOMs of the four radiologists are calculated and shown in Table III. We can observe that the FOMs of experts are higher than those of normal radiologists.

Table IV
Number of TP marks and FP marks under different threshold Tp

T_p	TP	FP
0.1	362	151
0.2	358	117
0.3	350	101
0.4	330	85
0.5	303	79
0.6	272	51
0.7	235	34
0.8	196	25
0.9	131	4

For the PCJO, we firstly study the influence of the probability threshold (T_p) by varying it from 0.1 to 0.9 with a step of 0.1. The numbers of TP and FP marks obtained by different T_p settings are shown in Table IV. We can see that when T_p is between 0.6 and 0.7, the TP and FP numbers of the PCJO seem close to those of radiologists.
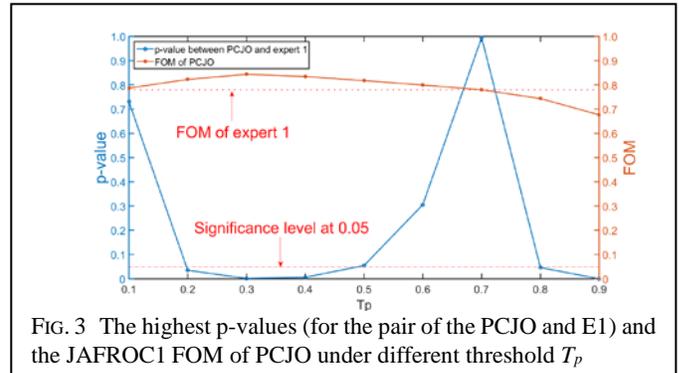
FIG. 3 The highest p-values (for the pair of the PCJO and E1) and the JAFROC1 FOM of PCJO under different threshold T_p

We then further find the relation between the *p*-value (for examining the JAFROC1 FOM difference between the PCJO and E1) and the probability threshold (T_p), as shown in Fig. 3. Note that for different thresholds, we tried different combinations of the number of steerable channel *K* and the number of scale-shiftable channel *J* to get the highest *p*-values. We see that when $T_p = 0.7$, there is no significant difference between the JAFROC1 FOM of the PCJO and that of E1. In order to make the JAFROC1 FOM of PCJO closest to that of E1, we retain the value 0.7 for the T_p for the abdominal ASiR LDCT images.

C. Discussion

We make a detailed survey of the JAFROC1 FOM and p-values between PCJO and radiologists. Results show that under certain VDP threshold and combinations of channels (K,J), there is no significant difference of performance between PCJO and radiologists.

We thus make a further investigation of the VDP threshold and the combination of channels. It is known that as the

difficulty increases, the HVS's detection probability threshold rises. We find that the threshold of FBP (0.8) and ASiR (0.7) are both high, which means lesions may be more difficult to be detected in LDCT images than that in normal-dose CT images.

We also investigate the relationship between the performance of FBP and that of ASiR. We see that there exists a behavior difference between FBP and ASiR. The average JAFROC1 FOM of ASiR is higher than that of FBP. Moreover, in FBP experiment, there is no significant difference between FOMs of each pair of radiologist as revealed in Table I, while in ASiR experiment, the FOMs of experts are comparable higher than those of normal radiologists, which can be proved in Table III and Fig. 3. We also find that the threshold is higher in the FBP experiment ($T_p = 0.8$) than that in the ASiR experiment ($T_p = 0.7$). What account for all these phenomena is that ASiR may have better image reconstruction performance than FBP does.

However, we can see that there exists several problems to be solved. In FBP experiment, the numbers of FP marks by experts are obviously higher than those by normal radiologists, as showed in Table I. In addition, in Table III, the number of TP marks by experts and that by normal radiologists are alike. These problems probably result from PCJO algorithm's adaptation to LDCT and subjective or objective factors that could influence the evaluation of radiologists.

V. CONCLUSIONS AND PERSPECTIVES

In this paper, we applied the extended PCJO to LDCT images. From what we discussed in the Chapter IV, we can finally draw following conclusions:

1. PCJO can approach human detection-localization task performance in the LDCT reconstruction image sets on the condition of the accurate VDP threshold T_p and the suitable combination of the number of steerable channels K and the number of scale-shiftable channels J;

2. ASiR has higher image reconstruction performance than FBP does.

The main contributions of this dissertation are: 1) an extensive application of PCJO on LDCT field, which reveals that PCJO can be adapt to LDCT image sets and shows its potential of auxiliary assessment ; 2) methods of validating the generality of the PCJO and related experimental values that can be provided for further study on different image sets and imaging modality; 3) an application of PCJO by comparing performance of different LDCT image reconstruction algorithms, which testifies that ASiR has better image reconstruction performance than FBP;

There exists two major limitations in the experiment: one is that the number of patients may be small because different radiologists may have different cognition of LDCT images; another limitation is that the scale of the image sets may be not large enough for training parameters. These two limitations require us to widen the scope of subjective evaluation and collect more accurate patients' data in our future experiments.

One of our perspectives that can be explored further is that applying the extended PCJO to LDCT image sets reconstructed by other algorithms, such as TV and Veo, to prove our conclusion. More imaging modalities, such as USI and SPECT, need to be investigated in the future as well. Another interesting aspect to be explored is that using 3D PCJO rather than 2D PCJO, because comparing to 2D data, 3D data is able to provide more analytical structure data and therefore can distinguish the noise and signal from the background easier.

REFERENCES

[1] Rusinek H, Naidich D P, McGuinness G, et al. Pulmonary nodule detection: low-dose versus conventional CT. Radiology, 209(1): 243-249, 1998

[2] Yang Chen, Xindao Yin, Luyao Shi, et al. Improving abdomen tumor low-dose CT images using a fast dictionary learning based processing, Physics in Medicine and Biology, 58(16):5803–5820, 2013.

[3] G. Mather, Foundations of Sensation and Perception. Hove and New York, USA: Psychology Press Ltd., Talor & Francis Inc., 2009.

[4] Barrett Harrison H., Myers Kyle J. Foundations of image science. John Wiley & Sons, 2013.

[5] He X., Park S. Model observers in medical imaging research. Theranostics. 2013.

[6] Zhang Lu, Goossens Bart, Cavaro-Menard Christine, et al. Channelized model observer for the detection and estimation of signals with unknown amplitude, orientation, and size. JOURNAL OF THE OPTICAL SOCIETY OF AMERICA A-OPTICS IMAGE SCIENCE AND VISION, 30(11): 2422-2432, 2013.

[7] Zhang Lu, Cavaro-Ménard Christine, Le Callet Patrick, et al. A perceptually relevant channelized joint observer (PCJO) for the detection-localization of parametric signals. IEEE transactions on medical imaging. 31(10): 1875-1888, 2012.

[8] S. Daly, The visible differences predictor: an algorithm for the assessment of image fidelity, Digital Image and Human Vision, A. B. Watson, Ed. Cambridge, MA: MIT Press, 1993.

[9] Zhang Lu. Numerical observers for the objective quality assessment of medical images. Université d'Angers, 2012.

[10] Goossens Bart. Multiresolution image models and estimation techniques. Ghent University, 2010.

[11] DP Chakraborty, KS Berbaum, Observer studies involving detection and localization: Modeling, analysis and validation. Medical Physics, 31(8): 2313-2330, 2004.

[12] Chakraborty Dev P. New developments in observer performance methodology in medical imaging. Seminars in nuclear medicine. Elsevier,401-418,2011.