A MULTI-SLICE MODEL OBSERVER FOR MEDICAL IMAGE QUALITY ASSESSMENT

L. Zhang [1], *C. Cavaro-Ménard* [2], *P. Le Callet* [3], *D. Ge* [4]

[1] IETR, CNRS UMR 6164, INSA Rennes, 20 avenue des Buttes de Coesmes, 35708 Rennes, France
[2] LISA, EA 4094, University of Angers, 62 avenue Notre Dame du Lac, 49000 Angers, France
[3] IRCCyN, UMR 6597, University of Nantes, Rue Christian Pauc, 44306 Nantes cedex 3, France
[4] LTSI, INSERM UMR 1099, Université de Rennes 1, Campus de Beaulieu, 35042 Rennes, France

ABSTRACT

Model observers (MOs) have been developed for the medical image quality assessment. Nowadays, numerous modern medical instruments are capable of producing 3D images, while few researchers have conducted MO studies on 3D data. In this paper, we propose a multi-slice MO when considering a relatively more realistic diagnostic task: the detection-localization of simulated multiple-sclerosis (MS) lesions on 3D magnetic resonance (MR) images. The jackknife free-response receiver operating characteristic (JAFROC) method was used to quantitatively analyse its performances and compare them with those of human observers. Our preliminary results showed that the proposed framework has the potential to approach human detection-localization task performance.

1. INTRODUCTION

Recently, the *task-based approach* [1] has been widely accepted for medical image quality assessment. In task-based evaluation, one or more observers perform one or more diagnostic tasks (e.g. detection task) using a set of images obtained from the imaging systems being evaluated. The imaging system that enables observers to get a better diagnostic task performance is said to be better. Since the evaluation using human observers is time-consuming and expensive, the evaluation using model observer (MO) has become an attractive alternative [2]. While numerous MOs have been proposed to perform a diagnostic task on solitary 2D images, those for 3D images have not been extensively studied yet. Compared to 2D images, 3D images provide more complete information which could allow for a better distinction between true abnormalities and noises/background structures. For some digital imaging modalities (e.g. MRI, CT), the 3D visualization of the image data is even becoming the standard for routine patient diagnostic care in different specialities (e.g. neurology, cardiology). Thus, more researchers have started to investigate 3D MOs which seems more clinically relevant.

Most existing 3D MOs are signal-known-exactly (SKE) MOs [3, 4, 5, 6, 7, 8] that can only perform the detection of one signal with known parameters (e.g. size and location) [9].

Although SKE MOs have successfully predicted human performance in SKE detection tasks [10], SKE is not clinically realistic. The truth is that human observers do not have such *a priori* knowledge in clinical practice. That is why there are increased demands of signal-known-statistically (SKS) MOs for which signal parameters are specified only statistically.

For the detection of a signal with variable location, scanning MOs have been proposed. They scan an image exhaustively, then choose the location that gives the largest test statistic (detection decision) as the tentative location and take the largest test statistic as the final test statistic. The idea has been adopted by Whitaker *et al.* [11] for 2D cases; and by Gifford *et al.* [12] for 3D cases, in which a two-phase visual search paradigm was proposed to improve the scanning process as a description of human search.

For the detection of a non-symmetrical signal with variable orientation, MOs could be generalized by using efficient channels which do not rely on signal symmetry, such as those proposed in [13, 14, 15]. These efficient channels are 2D, but they are not limited by the dimensionality of the problem by definition. An alternative (a simpler way) is to assume that the observer knows the mean reconstructed 3D signal profile to allow some signal variabilities in the test images [7].

In this paper, we propose an alternative to address the SKS task for 3D images. We tested the proposed MO for the detection-localization task of multiple non-symmetrical signals on MR images. Section 2 details the proposed MO. Section 3 presents the human observer study and shows the performance results of both observers. Section 4 concludes the paper and discusses possible future works.

2. PROPOSED MULTI-SLICE MODEL OBSERVER

2.1. 3D MS lesions simulation

The lesion is modelled by a 3D Gaussian function here:

$$[\boldsymbol{x}]_{\boldsymbol{p}} = a \exp\left(-\frac{1}{2}\left(\boldsymbol{p}-\boldsymbol{q}\right)^{\mathrm{t}} \mathbf{B}^{\mathrm{t}} \mathbf{D}^{-1} \mathbf{B}\left(\boldsymbol{p}-\boldsymbol{q}\right)\right), \quad (1)$$

where $[\boldsymbol{x}]_{\boldsymbol{p}}$ denotes the intensity value of the added signal at the 3-D coordinate $\boldsymbol{p}$. The signal intensity attenuation is

hereby modeled by a Gaussian function of peak amplitude a, centerd at $\boldsymbol{q}$. Without loss of generality, the diagonal matrix $\mathbf{D}$ specifies the equatorial radii σ and $b\sigma$ (along the x and y axes) and the polar radius $c\sigma$ (along the z-axis). We also used the x-y-z convention with independent Euler angles ϕ, θ, and ψ to construct the 3D rotation matrix $\mathbf{B}$: the rotations are applied firstly around the x-axis by ϕ, secondly around the y-axis by θ and thirdly around the z-axis by ψ; each basic rotation appears counter-clockwise and the axis about which it occurs points toward the observer.

An example of a set of original images and those with simulated 3D lesions is shown in Fig. 1. Note that a 3D signal appears brighter and possibly bigger on its central slice compared to its appearances on the adjacent slices.
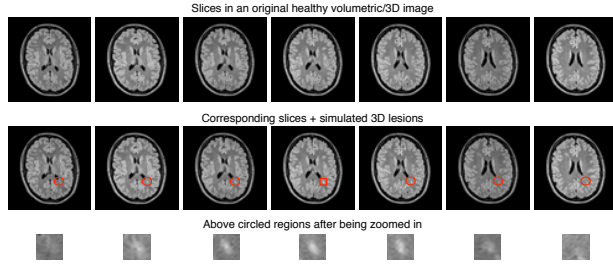

Slices in an original healthy volumetric/3D image

Corresponding slices + simulated 3D lesions

Above circled regions after being zoomed in

Fig. 1. The 1st row shows contiguous slices extracted from a healthy 3D image. The 2nd row shows the corresponding slices plus simulated 3D signals. One signal is highlighted to illustrate its appearances on 2D slices: marked by a red square on its central slice and by red circles on the adjacent slices. The highlighted regions are zoomed in the 3rd row.

2.2. msPCJO

The proposed MO is illustrated in Fig. 2, called multi-slice perceptually relevant channelized joint observer (msPCJO) hereafter. The msPCJO regards the detection-localization diagnostic task as two steps: (1) the first step is a global search to locate some 3D lesion candidates; (2) the second step is a cognitive analysis and interpretation of each candidate.

2.2.1. First step - Candidates localization

For this step, we take the assumption of conventional multi-slice MOs that radiologists interpret 3D images by pre-processing the image slice-after-slice (in x-y plane) and then integrate the information across the slices along z-axis [9].

We firstly apply a human visual system (HVS) model, visible differences predictor (VDP) [16], on each slice of an input volumetric image. The VDP produces a map of probabilities of detecting the differences between a *reference image* and a *distorted image* (as a function of their location in the input images), through several functions modelling various HVS properties (e.g. contrast sensitivity, orientation/frequency selectivity...). On the assumption that radiologists know what a healthy image should be like and refer to

it in their minds during the diagnostic process, we took an image of a healthy person as *reference image* (such as images of the 1st row in Fig. 1) and the same image with simulated signals as *distorted image* (such as images of the 2nd row in Fig. 1). Then the VDP actually predicted the probability that the signal was visible in the radiologists' perceptual domain.

With detection probability maps for all the slices, 3D coordinates (z-coordinate is the slice number) with detection probability superior to the empirical threshold T_p ($T_p = 0.9$ in this paper, since a detection probability often quoted is 90% [17] are preselected. These selected coordinates $\{\boldsymbol{y}_n\}_{1,...,N}$ are then classified based on a Gaussian mixture model (GMM) since the signal intensity is itself a Gaussian function (cf. Eq.(1)) and the problem here is to inverse the mixture (de-mix) of a linear superposition of an unknown number of Gaussian distributions. We here used one of the well-known methods of classification based on the GMM: the expectation-maximization (EM) algorithm [18]. For a given class number I, the EM algorithm draws confidence ellipsoids (centers and covariances). The Bayesian Information Criterion (BIC) is then computed to assess the number of classes. The minimization of the BIC solves the over-fitting problem (likelihood tends to infinity if one class is created for each data $\boldsymbol{x}_n$) by using a penalization term on the number of free parameters k (directly related to the number of classes):

$$\text{BIC} = -2\sum_{n=1}^{N} \log \left\{ \sum_{i=1}^{I} \pi_i \mathcal{N}(\boldsymbol{y}_n \mid \mu_i, \Sigma_i) \right\} + k\log(N)$$

The estimated number of classes is the estimated number of 3D signals and the corresponding centers of ellipsoids are then output as candidates centers.

The outputs of the first step are 3D coordinates of candidate centers and 2D blocks with a lesion candidate at the center extracted from the central slice of candidates. The block size $M \times M$ is pre-defined to cover biggest lesions.

2.2.2. Second step - CJO detection decision

In this step, a test statistic is calculated for each output 2D block using the channelized joint detection-estimation observer (CJO). The CJO was originally proposed for the detection of an SKS signal on a 2D image in our earlier study [15], that can be summarized into a training stage and a test stage:

Training stage: The inputs are two sets of image blocks (with-signal and without-signal) whose ground truth is known for the CJO, including the SKS signal profile x_α ($\boldsymbol{\alpha}$ is the signal parameter vector). Note that the signal is always located at the block center. The outputs of this stage are the estimated channelized reference signal $\boldsymbol{x}_0'$ and a template $\boldsymbol{w}$:

$$\widehat{\boldsymbol{x}_0'} = \langle \mathbf{A}_\alpha' \boldsymbol{x}_\alpha' \rangle = \langle \mathbf{A}_\alpha'(\mathbf{U}_0^{\mathrm{t}} \boldsymbol{x}_\alpha) \rangle, \qquad (2)$$

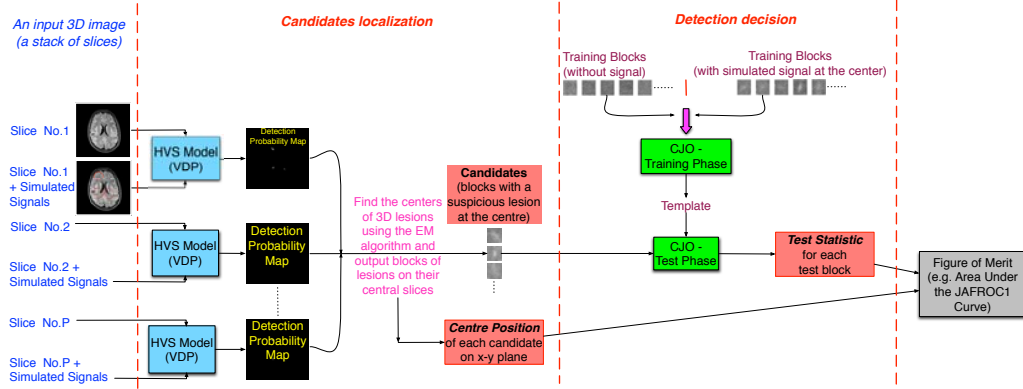$$\boldsymbol{w} = \left(\widehat{\Sigma_b'} \right)^{-1} \widehat{\boldsymbol{x}_0'}, \qquad (3)$$

where

Fig. 2. Structure diagram of the multi-slice PCJO (msPCJO): the outputs of each step are highlighted in red.

$$\widehat{\Sigma'_{b}} = \frac{1}{2}\langle (g' - \langle g' \mid H_0 \rangle)(g' - \langle g' \mid H_0 \rangle)^{\mathrm{t}} \mid H_0 \rangle +$$
$$\frac{1}{2}\langle [(g' - x'_\alpha) - \langle (g' - x'_\alpha) \mid H_1 \rangle]$$
$$\cdot [(g' - x'_\alpha) - \langle (g' - x'_\alpha) \mid H_1 \rangle]^{\mathrm{t}} \mid H_1 \rangle; \qquad (4)$$

$$g' = A'_\alpha U_0^{\mathrm{t}} g \qquad (5)$$

where the channel matrix U_0 reduces the data dimensionality, and A'_α maps the channelized parametric signal x'_α onto the channelized reference signal x'_0 : $A'_\alpha x'_\alpha = x'_0$.

To search the optimal parameters in the channel domain without loss of accuracy, one possible channel design (U_0 and A'_α) for the amplitude-orientation-scale-unknown case was proposed in [15] and adopted in this study. Note that K orientation-steerable channels and J scale-shiftable channels were used to construct U_0.

Test stage: w and $\widehat{x'_0}$ are used to estimate the signal parameters, as well as to calculate the test statistic λ for each input test image block whose ground truth is unknown to the CJO. For a fixed number of equally spaced signal parameters, the maximum of the test statistics yielded by all the possible parameter combinations is chosen as the final test statistic:

$$\lambda = \max_\alpha(\lambda_\alpha)$$
$$= \max_\alpha \left(\frac{w^{\mathrm{t}}}{\left\| U(A'_\alpha)^{\mathrm{t}} \right\|_F^2} \left(A'_\alpha g' - \frac{1}{2}\widehat{x'_0} \right) \right). \qquad (6)$$

Because the optimization in (6) takes place in the channel domain, the maximization of (6) is still relatively fast, even when done using a brute-force search.

3. RESULTS AND DISCUSSIONS

3.1. Experimental images

We collected 20 healthy subjects' MR brain images (3D axial stacks, T2 FLAIR sequence). Half of them were used for the CJO training stage, in which 1000 different 2D white matter (WM) regions were extracted (500 without signal and 500 with a simulated signal). The other 10 images were used as test images (for both human observers and the msPCJO), on which we added simulated 3D lesions as introduced in Section 2.2. The lesion locations were indicated manually to make sure that the lesions were added to the WM (quasi-randomly); the number of lesions in each 3D image was different, but fewer than 15 (to shorten the experiment time).
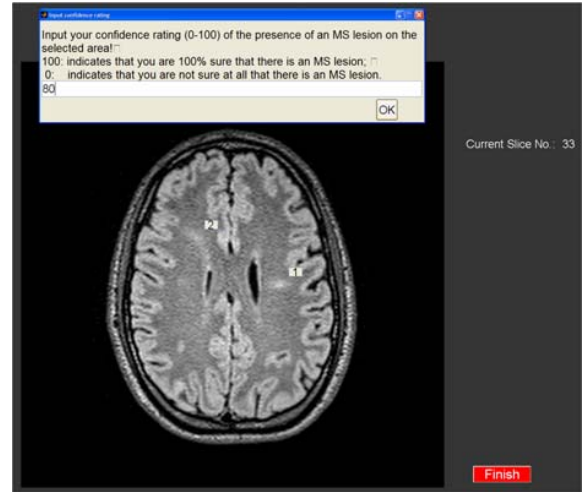

Fig. 3. Human observer study GUI.

3.2. Human observer study

We conducted a multi-slice free-response experiment, in which 6 radiologists were asked to detect and locate simulated MS lesions by scrolling back and forth the test images. When detecting a suspicious lesion, they selected its central slice and clicked the lesion center on the central slice. Then a dialog box popped up to ask the confidence rate (0-100).

Radiologists 1 and 2 were MS experts, with respectively 21 and 10 years of experience; Radiologists 3-6 were not MS experts, with respectively 6, 3, 8 and 5 years of experience.

3.3. Results and discussions

We used the software "JAFROC 4.1" [19] to evaluate the task performance, and conducted the DBM-MRMC significance test (integrated in the software) to examine JAFROC1 (figure of merite) values' difference between each pair of observers.

Radiologists' JAFROC1 values are shown in Table 1, p-values in Table 2. Table 3 shows the msPCJO's JAFROC1 values w.r.t. different number of steerable channels K and number of scale-shiftable channels J (cf. Section 2.2.2). Table 4 shows the p-values for the msPCJO with different (K, J) settings and 6 radiologists. Note that p-values< 0.05 are highlighted in red, which means that there is a significant difference between two corresponding JAFROC1 values.

Table 1. JAFROC1 values of six radiologists.

	JAFROC1 FOM	Standard Error
Radiologist 1	0.9267	0.0161
Radiologist 2	0.9106	0.0274
Radiologist 3	0.7837	0.0636
Radiologist 4	0.9192	0.0420
Radiologist 5	0.6547	0.0852
Radiologist 6	0.6320	0.0995

Table 2. The p-values for each pair of the radiologists.

	Rad1	Rad2	Rad3	Rad4	Rad5	Rad6
Rad 1	-	-	-	-	-	-
Rad 2	0.93	-	-	-	-	-
Rad 3	0.15	0.27	-	-	-	-
Rad 4	0.95	0.93	0.26	-	-	-
Rad 5	0.03	0.04	0.29	0.03	-	-
Rad 6	0.02	0.02	0.21	0.02	0.85	-

Table 3. The JAFROC1 values of the msPCJO.

(K,J)	JAFROC1 FOM	Standard Error
(3,3)	0.7076	0.1320
(3,4)	0.7645	0.1057
(3,5)	0.7820	0.0991
(5,3)	0.7029	0.1336
(5,4)	0.7622	0.1094
(5,5)	0.7785	0.1033
(7,3)	0.7099	0.1324
(7,4)	0.7622	0.1102
(7,5)	0.7797	0.1019

We see that from Table 3 that the number of steerable channels K hardly influences the msPCJO's performances, while the number of scale-shiftable channels J does. Table 4 shows that there is no significant difference between

Table 4. The p-values for the msPCJO vs. 6 radiologists.

(K,J)	Rad1	Rad2	Rad3	Rad4	Rad5	Rad6
(3,3)	0.02	0.02	0.40	0.02	0.56	0.40
(3,4)	0.07	0.09	0.83	0.09	0.22	0.14
(3,5)	0.11	0.14	0.98	0.13	0.16	0.10
(5,3)	0.01	0.02	0.37	0.02	0.59	0.43
(5,4)	0.07	0.10	0.81	0.08	0.23	0.15
(5,5)	0.10	0.13	0.95	0.12	0.17	0.11
(7,3)	0.02	0.04	0.41	0.02	0.54	0.39
(7,4)	0.07	0.10	0.81	0.08	0.23	0.15
(7,5)	0.10	0.15	0.96	0.12	0.17	0.10

the msPCJO's JAFROC1 values and those of 6 radiologists ($p > 0.05$) when $J > 3$.

Note that when $(K, J) = (3, 5)$ (i.e. $3 \times 5 = 15$ channels are used), the msPCJO attains its maximum detection-localization task performance (around 0.78), which is nearest to the 6 radiologists' average level 0.8045. That means when we use the msPCJO in practice, we can approach average human performance with not much calculation burden (less channels, less calculation burden).

4. CONCLUSION AND PERSPECTIVES

In this paper, we have proposed a novel 3D MO called the msPCJO in order to predict human performance in a study that has some relevance for diagnostic tasks which better describe clinical practice (multi-slice, multiple-signal...). To our knowledge, we are the first ones to propose such an MO that can perform the joint detection-localization of multiple 3D signals with random amplitude, orientation, size, location and number of lesions. By integrating an HVS model, the msPCJO also considers human visual characteristics. The inherent training stage in the 2nd step of the msPCJO allows modelling the training and the decision making process of radiologists to some extent.

We also conducted a human observer study close to the clinical paradigm to get the detection-localization task performance of six radiologists when volumetric information was available. The results are promising for our studied modality and pathology: the msPCJO is close to the average human performance with reasonable calculation amount.

One limitation of the study is the small number of 3D test images (mainly because it is difficult to get healthy patients). More test images will be needed to enrich the experimental validation in the further. Another interesting aspect to investigate consist in the extension of the 2D CJO to 3D CJO. Then in the 2nd step of the msPCJO, the CJO will be applied on 3D cubes with a lesion candidate at the center, instead of 2D blocks only on the central slice.

5. REFERENCES

[1] H.H. Barrett and K.J. Myerrs, *Foundations of Image Science*, John Wiley and Sons, Inc., Hoboken, New Jersey, USA, 2004.

[2] S. Leng, L. Yu, Y. Zhang, R. Carter, A.Y. Toledano, and C.H. McCollough, "Correlation between model observer and human observer performance in ct imaging when lesion location is uncertain," *Medical physics*, vol. 40, no. 8, pp. 081908, 2013.

[3] H. Liang, S. Park, B. D. Gallas, K. J. Myers, and A. Badano, "Image browsing in slow medical liquid crystal displays," *Academic Radiology*, vol. 15, pp. 370–382, 2008.

[4] J S. Kim, P. E. Kinahan, C. Lartizien, C. Comtat, and T. Lewellen, "A comparison of planar versus volumetric numerical observers for detection task performance in whole-body PET imaging," *Nuclear Science, IEEE Transactions on*, vol. 51, pp. 34–40, 2004.

[5] C. Lartizien, P. E. Kinahan, and C. Comtat, "Volumetric model and human observer comparisons of tumor detection for whole-body positron emission tomography," *Academic Radiology*, vol. 11, pp. 637–648, 2004.

[6] M. Chen, J. Bowsher, A. Baydush, K. Gilland, D. De-Long, and R. Jaszczak, "Using the hotelling observer on multislice and multi- view simulated SPECT myocardial images,," *Nuclear Science, IEEE Transactions on*, vol. 49, pp. 661–667, 2002.

[7] H.C. Gifford and M.A. King, "A comparison of human and model observers in multislice LROC studies," *IEEE Transactions on Medical Imaging*, vol. 24, no. 2, pp. 160–169, 2005.

[8] S. Young, S. Park, S. K. Anderson, A. Badano, K. J. Myers, and P. Bakic, "Estimating breast tomosynthesis performance in detection tasks with variable-background phantoms," in *Proceedings of SPIE*, 2009, vol. 7258, pp. 72580O.1–72580O.9.

[9] P. Ljiljana, G. Bart, V. Ewout, P. Subok, D.G. Brandon, B. Aldo, and P. Wilfried, "Channelized hotelling observers for the assessment of volumetric imaging data sets," *Journal of the Optical Society of America A*, vol. 28, no. 6, pp. 1145–1163, 2011.

[10] K.A. Craig and H.H. Barrett, "Human- and model-observer performance in ramp-spectrum noise: effects of regularization and object variability," *Journal of the Optical Society of America A*, vol. 18, pp. 473–488, 2001.

[11] M.K. Whitaker, E. Clarkson, and H.H. Barrett, "Estimating random signal parameters from noisy images with nuisance parameters: linear and scanning-linear methods," *Optics express*, vol. 16, no. 11, pp. 8150–8173, 2008.

[12] H.C. Gifford, "A visual-search model observer for multislice-multiview spect images," *Medical physics*, vol. 40, no. 9, pp. 092505, 2013.

[13] J.M. Witten, S. Park, and K.J. Myers, "Partial least squares: a method to estimate efficient channels for the ideal observers," *Medical Imaging, IEEE Transactions on*, vol. 29, no. 4, pp. 1050–1058, 2010.

[14] B. Goossens, L. Platisa, E. Vansteenkiste, and W. Philips, "The use of steerable channels for detecting asymmetrical signals with random orientations," in *Proceedings of SPIE*, 2010, pp. 76270S–76270S.

[15] L. Zhang, B. Goossens, C. Cavaro-Ménard, P. Le Callet, and Di Ge, "Channelized model observer for the detection and estimation of signals with unknown amplitude, orientation, and size," *Journal of the Optical Society of North America (JOSA) A*, vol. 30, pp. 2422 – 2432, 2013.

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

[17] G.A. Georgiou, "Probability of detection (pod) curves: derivation, applications and limitations," *Jacobi Consulting Limited Health and Safety Executive Research Report*, vol. 454, 2006.

[18] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *JSTOR*, vol. 39, no. 1, pp. 1–38, 1977.

[19] D. P. Chakraborty, "New developments in observer performance methodology in medical imaging," *Seminars in Nuclear Medicine*, vol. 41, pp. 401–418, 2011.