Disponible en ligne sur
ScienceDirect
www.sciencedirect.com

Elsevier Masson France
EM|consulte
www.em-consulte.com

IRBM

IRBM 35 (2014) 214–224

General review

An overview of model observers

L. Zhang [a,*], C. Cavaro-Ménard [b], P. Le Callet [c]

[a] IETR-Image, CNRS UMR 6164, INSA Rennes, 20, avenue des Buttes-de-Coesmes, CS 70839, 35708 Rennes, France
[b] LISA, EA 4094, University of Angers, 62, avenue Notre-Dame-du-Lac, 49000 Angers, France
[c] IRCCyN, UMR-6597, University of Nantes, rue Christian-Pauc, BP 50609, 44306 Nantes cedex 3, France

Received 25 October 2013; received in revised form 21 February 2014; accepted 7 April 2014
Available online 2 June 2014

Abstract

Nowadays, model observers have been used more and more for the objective quality assessment of medical images. Model observers have been developed from signal known exactly (SKE) task to signal known statistically (SKS) task, from single-slice (2D) to multi-slice (3D), in order to be more clinical relevant. In this paper, we give an overview of existing model observers up to date.

Keywords: Model observer; Medical image quality assessment; HVS

1. Introduction

Nowadays, with the development of medical imaging techniques, medical experts (e.g. radiologists and physicians) face numerous choices among different image acquisition systems, image post-processing algorithms (such as compression and watermarking) and display systems. There is thus an urgent demand for a good image quality assessment approach, from the standpoint of end-users (medical experts).

Development and design of medical imaging technologies should take into account performance of radiologists on relevant tasks as this has an important impact on patient care. Towards this purpose, the quality of the medical image should be quantified by its effectiveness with respect to its intended purpose, which can be represented by the diagnostic task performance. This type of approach is called task-based approach [1]. In task-based evaluation one or more observers performs one or more tasks using a set of images obtained from the imaging system being evaluated. Then the system/algorithm that enables observers to yield a better task performance is considered to be better.

Using human observers has several limitations, e.g. time-consuming, costly, variances existing between and within human observers' responses [2]. Consequently, model observers (MOs) appeared as a promising surrogate for human observers in the field of medical image quality assessment.

In this paper, we give an overview of current MOs, from signal known exactly (SKE) task to signal known statistically (SKS) task, from single-slice (2D) to multi-slice (3D). The following of the paper is organized as follows: Section 2 introduces the mathematical backgrounds of MOs (e.g. background and signal model). Section 4 presents the most popular MOs for the SKE task. Section 5 summaries the existing MOs for the SKS task, including two new MOs proposed in our previous works. Section 6 presents the multi-slice MOs, including one new MO proposed in our previous work. A conclusion is given in Section 7.

2. Background and signal models

In general, MO is designed to detect a signal on a noisy background. The problem can be seen as the validation of one of two exclusive hypotheses formulated as follows:

$$\mathcal{H}_h : g = hx + b, h = 0, 1 \quad (1)$$

where g is an M * 1 column vector representing the observed image consisting of M pixels; x denotes the signal and b denotes the noisy background; the absence/presence of the signal is controlled by the binary variable h. Note that normally the amount

* Corresponding author.
E-mail address: lu.ge@insa-rennes.fr (L. Zhang).

http://dx.doi.org/10.1016/j.irbm.2014.04.002
1959-0318/

of noise is assumed to be so small that it does not disturb the statistical properties of the background.

In the literature, the background b and the signal x in Eq. [1] are usually synthetized (thus their characteristics can be totally controlled) for mathematical modeling and traceability.

Four common background models are illustrated on Fig. 1: white gaussian background (WNB), correlated gaussian background (CGB), lumpy background (LB) [3] and clustered lumpy background (CLB) [4]. A simulated signal can then be added to a background. Examples are shown on Fig. 2, where the signal is modeled by a 2D elliptical Gaussian function and added at the center of backgrounds.

A general form of the signal is:

$$[x]_p = f_\alpha(p) \tag{2}$$

where p denotes 2D coordinates on the background, and $f_\alpha(p)$ is a function that represents a general parametrized signal with parameter α:

$$\alpha = \left[a, \theta, b, \sigma, q\right] \tag{3}$$

where a represents the signal amplitude (intensity), θ represents the signal orientation, b represents the signal shape, σ represents the signal scale, and q represents the signal center position. If all the signal parameters in α are known *a priori* by observers, the detection task is a SKE task. Otherwise, if at least one of the signal parameters is specified by a probability density function (PDF), it is a SKS task [5].

3. Principle and evaluation of model observers

For all MOs, a scalar test statistic $\lambda(g)$ is computed via a discriminant function of the image g. Then, a decision is made in favor of the hypothesis $\mathcal{H}_1$ if the test statistic $\lambda(g)$ is greater than a decision criterion λ_c; otherwise $\mathcal{H}_0$ is selected. The decision rule can be represented as:

$$\lambda(g) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda_c \tag{4}$$

MOs differ by their discriminant functions.

In order to quantify and characterize the performances of MOs, different figures-of-merit (FOMs) have been proposed for different diagnostic tasks. In general, a higher value of the FOM means a better task performance.

Given the simplest diagnostic task – the detection task of one signal on an image, we can use the receiver operating characteristic (ROC) curve to depict the MO's performance. A ROC curve is a graphical plot of the fraction of true positives out of the total actual positives vs. the fraction of false positives out of the total actual negatives, at various decision criterion settings [6]. One ROC-based FOM is the area under the ROC curve (AUC), which is equal to the probability that a MO will rank a randomly chosen positive image higher than a randomly chosen negative one [7]. In practice, the AUC could be calculated by using an average of a number of trapezoidal approximations [8].

Another FOM for this task is the detectability index, which is also called signal-to-noise ratio (SNR) in some papers.

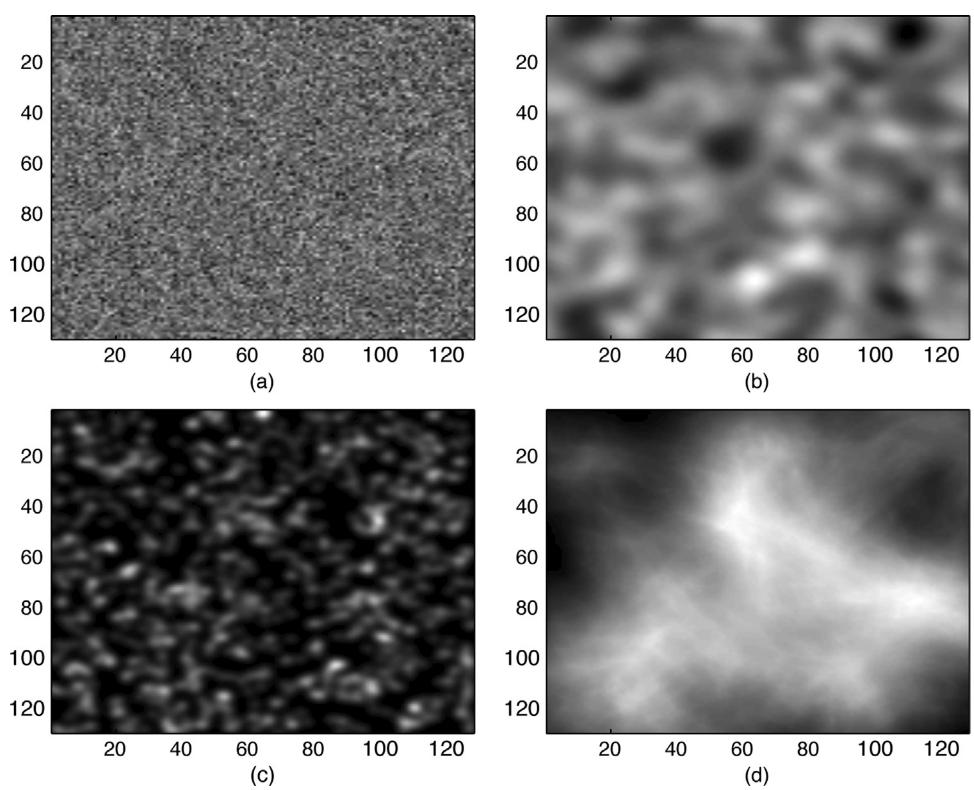

Fig. 1. Examples of four background models: a: white gaussian background (WNB); b: correlated gaussian background (CGB); c: lumpy background (LB); d: clustered lumpy background (CLB).