Physica Medica xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Physica Medica

journal homepage: http://www.physicamedica.com

Original paper

Can the channelized Hotelling observer including aspects of the human visual system predict human observer performance in mammography?

R.W. Bouwman [a,*], M. Goffi [b], R.E. van Engen [a], M.J.M Broeders [a,c], D.R. Dance [d,e], K.C. Young [d,e], W.J.H. Veldkamp [a,f]

[a] Dutch Reference Centre for Screening (LRCB), Radboud University Medical Centre, The Netherlands
[b] Department of Medical Physics, Elisabeth TweeSteden ziekenhuis, The Netherlands
[c] Radboud Institute for Health Sciences (RIHS), Radboud University Medical Centre, The Netherlands
[d] National Co-ordinating Centre for the Physics of Mammography (NCCPM), Royal Surrey County Hospital, United Kingdom
[e] Department of Physics, University of Surrey, United Kingdom
[f] Department of Radiology, Leiden University Medical Centre (LUMC), The Netherlands

ARTICLE INFO

Article history:
Received 29 September 2016
Received in Revised form 12 December 2016
Accepted 21 December 2016
Available online xxxx

Keywords:
Model observer
Mammography
Clinical image quality
Human visual system

ABSTRACT

Purpose: In mammography, images are processed prior to display. Model observers (MO) are candidates to objectively evaluate processed images if they can predict human observer performance for detail detection. The aim of this study was to investigate if the channelized Hotelling observer (CHO) can be configured to predict human observer performance in mammography like images.
Methods: The performance correlation between human observers and CHO has been evaluated using different channel-sets and by including aspects of the human visual system (HVS). The correlation was investigated for the detection of disk-shaped details in simulated white noise (WN) and clustered lumpy backgrounds (CLB) images, representing respectively quantum noise limited and mammography like images. The images were scored by the MO and five human observers in 2-alternative forced choice experiments.
Results: For WN images the most useful formulation of the CHO to predict human observer performance was obtained using three difference of Gaussian channels without adding HVS aspects ($R^2_{LR} = 0.62$). For CLB images the most useful formulation was the partial least square channel-set without adding HVS aspects ($R^2_{LR} = 0.71$). The correlation was affected by detail size and background.
Conclusions: This study has shown that the CHO can predict human observer performance. Due to object size and background dependency it is important that the range of object sizes and allowed variability in background are specified and validated carefully before the CHO can be implemented for objective image quality assessment.

1. Introduction

Current methodologies to evaluate physical image quality (PIQ) in digital mammography (DM) use unprocessed images of simple homogeneous phantoms [1] and do not evaluate the processing which is applied to clinical images [2]. This means that the quality of the images as seen by radiologists may be different from those used in PIQ evaluations. The kind and amount of processing differs between manufacturers and software versions. Several studies have shown that image processing affects cancer detection [3,4] and emphasize the need for objective measurements of image quality as perceived by radiologists which is referred to as clinical image quality (CIQ). To allow evaluation of CIQ, a new method has to be developed, which incorporates clinically relevant structures and does not make assumptions about linearity of the images. Statistical linear model observers (MO) have been suggested for this purpose [5], but before MOs can be introduced for CIQ assessment the correlation between human and MOs must be known and their use should be validated.

MOs can be used in for example detection experiments. Among the several existing MOs, the non-pre-whitening (NPW) MO and the channelized Hotelling observer (CHO) are candidates for the evaluation of CIQ. The NPW MO estimates performance based on the correlation between the expected signal and the images. For this MO the signal is assumed to be known exactly whereas the

* Corresponding author at: Dutch Reference Centre for Screening (LRCB), Radboud University Medical Centre, PO Box 6873, 6503 GJ Nijmegen, The Netherlands.

E-mail address: r.bouwman@lrcb.nl (R.W. Bouwman).

http://dx.doi.org/10.1016/j.ejmp.2016.12.015
1120-1797/

observer template from the CHO is based on the channelized covariance of the images [6]. In a previous paper [2], the NPW MO was evaluated to predict human observer performance by including aspects of the human visual system (HVS). For the NPW, it was found that the correlation between human and model performance had a large dependency with the object to be detected and the background structure used. This means that the correlation can only be used for image quality purposes in a limited range of object size and background variability, which has to be specified and validated carefully. In this study, the CHO has been evaluated in a similar way to decide the appropriateness of the CHO for CIQ assessment.

In this study, which is an extension of previous work [2,7], the most commonly used channel-sets in combination with different aspects of the human visual system (HVS) [8,9] have been studied. For this purpose we evaluated the correlation between human and MO performance for a simple detection task of a signal with various sizes, in simulated white noise (WN) and clustered lumpy backgrounds (CLB, [10]). These types of images were chosen to represent images from a quantum noise limited ideal system and images with clinically realistic structures. The following aspects of the HVS were taken into account in the CHO: 1) different eye filters [2], 2) contrast masking of the signal to be detected due to fluctuations in the background luminance (noise mask, NM [8]) and 3) the addition of a psychometric function (probability map, PM [9]) to account for the human decision making process. Since different formulations of the CHO are used throughout this manuscript, we use CHO as a generic term for all the formulations considered.

2. Materials and methods

2.1. Channelized Hotelling observer

In detection experiments the observer has to decide on the presence or absence of a signal (for example a mass) in images by dividing them into two classes: signal present (class 1) and signal absent (class 2). The goal is to determine how well both classes can be distinguished which is expressed in terms of the detectability index, d' [6]. For MOs, d' is estimated from the decision variables (λ) assigned to each image. For linear MOs, λ is described by a linear transformation between an observer template ($\boldsymbol{w}$) and the image vector ($\boldsymbol{g_{i,j}}$):

$$\lambda_{i,j} = \mathbf{w}^t\mathbf{g}_{i,j} + \varepsilon \tag{1}$$

where i represents the image class (1,2), j the image number, t the transpose matrix, and ε represents a model of internal noise to account for inconsistencies by human observers [11]. The inclusion of internal noise is beyond the scope of this paper and will not be discussed further. To reduce the computational complexity the two-dimensional image is written as a one-dimensional vector (bold symbols).

The linear model observer that maximizes the decision variables signal to noise ratio,, under assumptions of Gaussian image statistics, is the Hotelling observer, whose template, w_{HO}, is estimated using Eq. (2) [6]:

$$\mathbf{w}_{HO} = \mathbf{s}^t\boldsymbol{\kappa}^{-1} \tag{2}$$

where $\mathbf{s}$ is the difference between the mean data of both classes and $\boldsymbol{\kappa}$ is the covariance matrix of the images. Due to the high dimensionality of $\boldsymbol{\kappa}$, inversion is computationally demanding. The dimensionality can be reduced by channelizing the images, and the MO is then referred to as the channelized Hotelling observer (CHO). The formulation of the observer template for the CHO, $\mathbf{w}_{CHO,}$ is:

$$\mathbf{w}_{CHO} = \mathbf{K}_c^{-1}[\overline{\mathbf{g}_{c1}} - \overline{\mathbf{g}_{c2}}] \tag{3}$$

where $\overline{\mathbf{g}_{ci}}$ is the mean channelized image vector of class i and $\mathbf{K}_c$ the interclass channel covariance matrix estimated using:

$$\mathbf{K}_c = \frac{1}{2}[\kappa_{c1} + \kappa_{c2}] \tag{4}$$

with $\boldsymbol{\kappa}_{ci}$:

$$\boldsymbol{\kappa}_{ci} = cov(U^t g_i) \tag{5}$$

where U^t is the transpose of the channel matrix and $\mathbf{g_i}$ the image vector of all images of class i. The process of estimating the CHO decision variables consists of two steps. In the first step the template, $\mathbf{w}_{CHO}$, is determined, which is referred to as "training". The second step is the estimation of the decision variables, referred to as "testing".

Several channel-sets have been proposed to reduce the dimensionality of the covariance matrix. These channel-sets could either be anthropomorphic or efficient [12]. Anthropomorphic channels aim to include HVS aspects, while efficient channels aim to process all available information in such a way that the performance is comparable to the ideal linear MO (the HO). Depending on the objective of the experiment, the desired channel-set is chosen. Additional constraints to the channel-sets can be formulated based on prior knowledge of the task. For example, if the signal to be detected and the noise power spectrum (NPS) of the background structure are rotationally symmetric, the channels chosen could also be rotationally symmetric. In this study 5 different channel-sets in combination with aspects of the HVS have been evaluated: three anthropomorphic channel-sets: Gabor, difference of Gaussian (DoG) and differences of mesa (DoM) and two efficient channel-sets: the Laguerre-Gauss (LG) and the partial least square (PLS) channel-set. Details of each channel-set are given below and are summarized in Table 1 and Fig. 1.

2.1.1. Gabor

Gabor channels have been used to model the response of cells in the visual cortex for pattern recognition and are given by [13,14]:

$$C_{Gabor}(x,y) = e^{\left[\frac{-4\ln 2((x-x_0)^2+(y-y_0)^2)}{W_s^2}\right]}\cos[2\pi f_c((x-x_0)cos\theta + (y-y_0)sin\theta) + \beta] \tag{6}$$

Watson [15] argued that each channel should have a bandwidth (W_s in cycles per pixel (cpp)) of about 1 octave, 8 basic filters should be used with centre frequencies (f_c) in the range of 0.25–32 cycles per degree (cpd), each with two phases (β) to preserve the odd (0) and even ($\pi/2$) information in the images and 5 orientations (θ): 0, $2\pi/5$, $4\pi/5$, $6\pi/5$ and $8\pi/5$. Thus, the Gabor channel-set will then consist of 80 (8x5x2) channels centred around the midpoint of the disk, x_0 and y_0.

Table 1
Details of the channel-sets used.

Channel-set	Number of channels	Type of channel	Features
Gabor	80	Anthropomorphic	Frequency, orientation and phase
DoG	3,10	Anthropomorphic	Frequency
DoM	31	Anthropomorphic	Frequency and orientation
LG	Optimized*	Efficient	Frequency and optional orientation
PLS	Optimized*	Efficient	Fully based on the image set.

*The number of channels is optimized based on the training set to maximize the decision variable.