Accounting for anatomical noise in tumor localization tasks with visual-search and channelized Hotelling observers

Anando Sen and Howard C. Gifford

Department of Biomedical Engineering, University of Houston, Houston, Texas, USA

Abstract

Model observers are used for image quality assessment in imaging system designs. Models observers that can mimic human perceptions are of particular interest. In development of model observers it is important to account for the effects of quantum and anatomical noise. In this study, we present the comparison between the channelized Hotelling observer (CHO) and the visual-search (VS) observer as two ways of accounting for anatomical noise. The application considered is tumor localization in prostate SPECT. The SPECT acquisition is simulated through an analytic projector. The projections are reconstructed using the OSEM algorithm. Observer studies are carried out with the CHO, various forms of the VS observer and human observers. Observer performance is assessed by localization ROC (LROC) analysis. Results indicate that human observer correlates better with the VS observer as compared to the CHO.

Introduction

Tomographic imaging is routinely used for cancer diagnostics. Medical images are read by radiologists, who make the final judgment about the presence of cancer. Image quality is crucial for the radiologist to make the best possible decision However human studies at every stage of developmental research is time consuming. This has led to the development of **model observers which are mathematical models for detection and localization of tumors** in medical images. There is particular interest in developing model observers that can mimic human performance. When developing model observers, it is essential to account for the sources of variations in images. The two major sources are **quantum and anatomical noise**. While quantum noise is a result of random photon distributions, anatomical noise is caused by the variations in patient anatomy.

Scanning Channelized Hotelling Observer

- Typical **scanning observer** which computes perception measurements at every location within a region of interest.
- Accounts for quantum and anatomical variations in images through **location-specific class covariance matrices**.
- Heavily **dependent on prior information** about the background and the signal
- Requires **extensive preprocessing** (the covariance matrices computed ahead of time). Heavier when anatomical variations are included.
- Debatable if pre-whitening is equivalent to a human's response to anatomical noise.

Visual Search Observers

- Two step process **based on the VS paradigm for a radiologist's interpretation of images**. First step involves a search for suspicious tumor locations which are then analyzed in detail.
- Both search and analysis steps are **characterized by image features**.
- In **agreement with radiologists** who focus on a few areas of the image rather than scan every point.
- No requirement for covariance matrices. Search step **implicitly accounts for anatomical noise**.
- Previously Hotelling-type discriminants have been used for analysis. Hence VS observer is still tied to prior information.

In this poster we present a comparison of the CHO and the VS observers. We discuss various versions of the VS observer including an **adaptive VS (AVS) observer.** The AVS observer uses a **background-independent linear discriminant** for the analysis stage. This discriminant is computed by **training on several image features** and selecting the best possible feature combinations. The application considered is tumor localization in prostate SPECT.

Model Observers

For single-target LROC studies, an observer selects the most likely tumor location in each image and rates it for the presence or absence of a tumor. For typical scanning observers these are obtained by **the maximization of a location-specific perception measurement** λ_j within the region of interest Ω.

Rating: $\lambda = \max_{j\in\Omega} \lambda_j$

Location: $r = \text{argmax}_{j\in\Omega}\lambda_j$

Channelized Hotelling Observer

- Location-specific covariance matrices (for all $j \in \Omega$) defined by

$$\boldsymbol{K}_j = \boldsymbol{U}_j^t \boldsymbol{K} \boldsymbol{U}_j$$

where $\boldsymbol{K}$ is the ensemble class covariance matrix accounting for all quantum and anatomical variations, $\boldsymbol{U}$ is the channel matrix whose columns are the frequency channels $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_C\}$. The subscript indicates the shift of the channel response to the jth location. We have used a set of three difference of Gaussian channels.

- For computational efficiency the covariance is decomposed into quantum and anatomical components and added

$$\boldsymbol{K}_j = \boldsymbol{K}_j^{quant} + \boldsymbol{K}_j^{anat}$$

- The CHO observer template at the jth location is given by

$$\boldsymbol{w}_j = \boldsymbol{U}_j \boldsymbol{K}_j^{-1} \boldsymbol{U}_j^t \bar{\boldsymbol{s}}_j$$

with $\bar{s}$ denoting the mean target profile and the subscript indicating its shift to the jth location.

- The CHO perception measurements for a test image $\boldsymbol{f}$ are given by

$$\lambda_j = \boldsymbol{w}_j^T (\boldsymbol{f} - \boldsymbol{b})$$

where $\boldsymbol{b}$ is the mean background over all quantum and anatomical realizations.

Visual Search Observers with Hotelling-type Discriminants

- Target-dependent feature images $\boldsymbol{Z}^{\mathcal{O}}$ defined by the cross-correlation

$$\boldsymbol{Z}^{\mathcal{O}} = \mathcal{O}(\bar{\boldsymbol{s}}) \star\star \mathcal{O}(\boldsymbol{f})$$

- Gradient-based feature is used for the search. $\boldsymbol{Z}^{\nabla}$ is segmented into blobs $B_1, B_2, \dots$ by applying the **watershed algorithm**. The blob maxima within Ω constitute the reduced set Ω' for further analysis

$$\Omega' = \left\{k \in \Omega : k = \text{argmax}_{j\in B_i} z_j^{\nabla} \text{ for some } B_i\right\}$$

- Analysis performed by the channel feature $\boldsymbol{Z}^{\boldsymbol{U}}$ for the $\boldsymbol{U}$ defined above. This is equivalent to a non-prewhitening approximation of the CHO i.e. the channelized non-prewhitening observer

$$\boldsymbol{\lambda}_j = [\boldsymbol{U}_j \boldsymbol{U}_j^t \bar{\boldsymbol{s}}_j]^T (\boldsymbol{f} - \boldsymbol{b})$$

- **Possible for the VS observer to miss the actual tumor location** during the search. This is an important property for modelling human perceptions as it is possible for humans to miss the tumor location.

An Adaptive Visual-Search Observer

- Four target-based features used: the intensity I, the gradient ∇, the Laplacian Δ and the channel operator $\boldsymbol{U}$.
- **Adaptive feature selection using a set of training images**:
  1. Perform the search as described for the VS observer. Let Γ be the union of the sets Ω' for each image.
  2. Locations within Γ are **denoted by a four-element feature vector**. The features are standardized using global means and standard deviations. Denote the location vectors with standardized features by l.
  3. Compute the **pointwise means $\boldsymbol{\mu_1}$ and $\boldsymbol{\mu_0}$** of the positive and the negative locations. Let
  $$\delta = \mu_1 - \mu_0.$$
  4. For 4 features we have 15 feature combinations. Denote a feature combination by a 4-element indicator vector $\xi_i, i = 1, \dots, 15$ with each element of ξ_i indicating whether a feature is included in the feature set. For a location l and a feature combination ξ_i the analysis statistic is given by
  $$\lambda_j = (\delta . \xi_i) . l$$
  5. The feature combinations **maximizing the fraction of correct localization** will be used in the testing stage.
- The **testing stage** requires the four element vector δ, the maximizing feature combinations (there may be multiple) as well as global means and standard deviations of each feature from the training stage
  1. For test images apply the single feature search as for the VS observer
  2. Perform the **analysis using each of the maximizing feature combinations** as described in the training.
  3. Compute the LROC area for each of the maximizing feature combinations and use their average as the final figure of merit.
- The analysis stage is now **independent of the background**.

SPECT Simulation and Observer Studies

Clinically realistic **In-111 uptake ratios** are assigned to the organs in an **XCAT phantom**. Tumor-present cases contain **a single spherical, soft-tissue target** of diameter 1 cm in the **prostate or the lymph nodes**. An **analytic projector** is used for the SPECT acquisition. Poisson noise with **4.1 to 5,6 million counts** is added. The projections are reconstructed using the **OSEM algorithm** with 2 and 6 iterations and different post-smoothing blur widths (FWHM between 0 and 4 pixels) are applied. 2D slices through lesion centers are extracted for observer studies.

Observer studies are conducted with 150 images per strategy: 50 training and 100 test. Each strategy is defined by number of iterations and Gaussian blur width. The CHO, the VS observer (with and without background subtraction [NB] for the analysis), the AVS and a set of four human observers are used for the observer studies.

Results

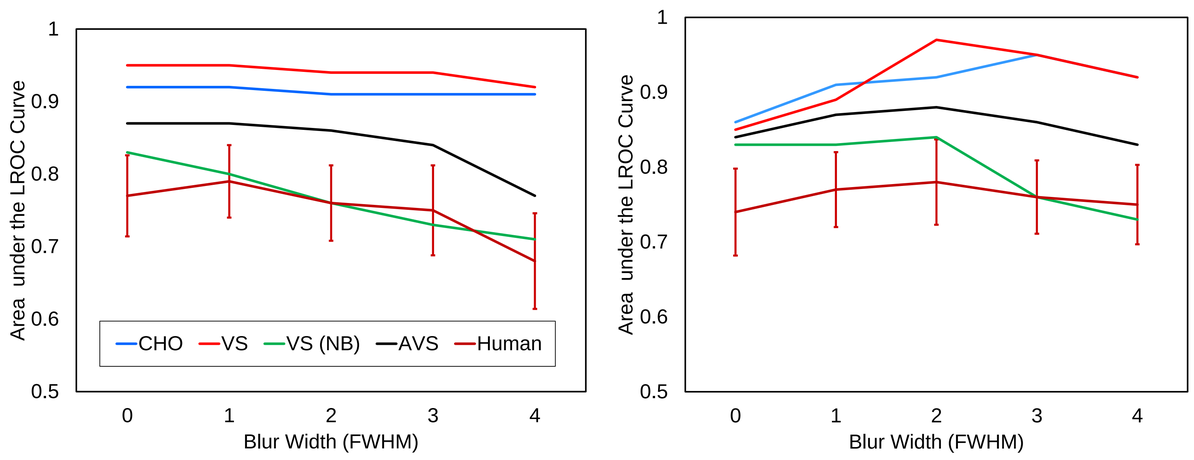

Figure: The observer performance (LROC area) as a function of the blur width (FWHM) for 2 iterations (left) and 6 iterations (right).

- Both the CHO and the VS observer **outperform the human observers.**
- The CHO and the VS observer yield similar performances though the VS observer **follows human trends better**.
- The AVS **lowers the gap** between the models and the humans and provides the **best correlation** with humans.
- Though the VS (NB) is closest to the humans in terms of performance the **trends are unstable**. This leads to **poor correlation** coefficients (evident for the 6 iterations case).

Observer	2 iterations	6 iterations
CHO	0.65	0.53
VS	0.97	0.78
VS(NB)	0.79	0.38
AVS	0.98	0.91

Table: Correlation coefficients of the model observers with the human observer for 2 and 6 iterations

Discussion and Conclusions

While the CHO uses image class statistics to account for anatomical noise, the VS observers respond to anatomical noise on **an image-by-image basis** which is in agreement with humans. The possible tumor candidates selected by both humans and the VS observers **could be due to high uptake (actual tumor location) or anatomical/quantum blobs.** Hence the VS observer is a **more accurate model** for human perception.

Despite heavy pre-processing (approximately 2 minutes) the CHO is **substantially slower than the VS observer**. For our study sets the application of the CHO and VS observers take about 2 minutes and 5 seconds respectively. Due to the training requirements, the **AVS observer is somewhat slower – about a minute – but still faster than the CHO**. Uncertainty in anatomical structures would further complicate the task for the CHO.

The AVS observer is an important step towards **eliminating prior knowledge** in detection-localization tasks. The use of Hotelling-type discriminants for analysis in the VS observer leads to **overestimates of human performance** as much as with the CHO. The AVS observer reduces the model-human performance gap and also has excellent correlation with the humans. Further enhancement of the AVS will include a location-based adaptation.

We have compared the CHO and the VS observer as two ways of accounting for anatomical noise in tumor localization tasks. The VS observer is a **more accurate model for humans and has the added advantages of superior correlation and computational efficiency**. This comparison of model observers argues for continued development of the VS observer especially in the adaptive format.

This research is supported by the National Institute of Biomedical Imaging and Bioengineering (NIBIB) grant R01EB012070