Questions and Answers

Chapter 13

Questions

(Q1) *What is the role of CADetection and CADiagnosis techniques in breast cancer imaging and mammography, in particular?*

(Q2) *What are the basic elements and characteristics of a CADiagnosis scheme?*

(Q3) *What is the clinical "visual analysis" system of mammographic calcifications and how can it be used for feature selection in a CADiagnosis algorithm?*

(Q4) *What property exists between wavelet expansion images that does not exist in an arbitrary filter bank output?*

(Q5) *What can you say about a time series signal that is nothing but a spike a* $t = t_0$ *with respect to its Fourier properties?*

(Q6) *An image with long-range positive correlation will have large Fourier components in what part of the frequency domain?*

(Q7) *Explain what a band pass filter is and what it may be used for.*

(Q8) *What is white noise and give an example of it from common observation?*

(Q9) *Given a low frequency signal of interest buried in white noise, what kind of filter would work for lessening the influence of the noise?*

(Q10) *What are the validation steps of a CADiagnosis scheme?*

Answers

(A1) CADetection algorithms for mammography are computer methodologies developed to assist the radiologist in the interpretation of a mammogram and specifically in the identification of suspicious areas. CADetection algorithms play the role of “second reader” by marking potential abnormalities on the image and, thus, improving detection sensitivity. CADiagnosis algorithms are also computer methodologies developed to assist the radiologist in the differentiation between benign and malignant lesions detected on mammograms. CADiagnosis algorithms play the role of a “second opinion” by providing a diagnostic decision for a finding often in addition to an automated detection outcome. The goal of CADiagnosis is to improve patient management and follow-up.

(A2) CADiagnosis algorithms characterize *detected* lesions as benign or malignant, thus, fusing the processes of detection and diagnosis. Conceptually a CADiagnosis system may be viewed as consisting of two main blocks, one for *detection* and one for *classification*. The detection block can be either a human observer or a CADetection system. In the former case, the radiologist identifies suspicious lesions and provides the characteristics (features) to the classification block of the algorithm. In the latter case, pre-processing, detection, and segmentation stages are usually parts of the detection block while feature extraction and classification are parts of the classification block. Statistical and non-statistical classifiers may be used for classification, the ranked ordered output of which is usually transformed to correspond to “likelihood of malignancy” for the lesion of interest.

(A3) Radiologists have attempted to differentiate between benign and malignant calcifications based on visual assessment of their shape, morphology, and distribution. By correlating these characteristics with pathology, a clinical visual analysis system was developed that has shown to yield a sensitivity of 97.6% (correct identification of cancers associated with calcifications) and a specificity of 73.3% (correct identification of benign cases associated with calcifications). This analysis system was used in the development of the Breast Imaging Reporting and Data System, the standard recommended by the American College of Radiology for mammography interpretation.

(A4) Wavelet expansion images are linearly independent. That is, the inner product between any two image arrays always vanishes. Likewise, the entire collection is complete indicating that the sum returns the image exactly.

(A5) The power spectrum of this "spike" signal will be flat. That is, all frequencies will be represented equally.

(A6) An image with long-range positive correlation is predominantly a low frequency signal, i.e., it has disproportionately high amplitude components in the low frequency domain.

(A7) Band pass filter is a mechanism used to block unwanted frequency components. It is implemented either as convolution (a weighted average of sorts) in the signal domain or multiplication in the frequency domain.

(A8) White noise is noise produced by combining signals (sounds) of all different frequencies together. It consists of a continuous and uniform frequency spectrum, i.e., frequencies are distributed equally over the range of a specified frequency band. The sound of ocean waves or waterfalls are good examples of white noise. Likewise, when tuning a television to a channel where there is no signal also provides an example of white noise or snow.

(A9) A low pass filter is best suited to reduce the influence of white noise in this scenario.

(A10) The validation of a CADiagnosis algorithm should theoretically follow the six-level model proposed by Fryback and Thornbury (Med. Decis. Making Vol. 11, No. 2, pp. 88–94, 1991). In practice, the first three levels of the model are usually sufficient for CADiagnosis applications that involve laboratory and initial clinical testing. Laboratory training and testing of a

CADiagnosis algorithm requires a series of well-designed tests to avoid biased results particularly when small datasets are available and/or large feature sets are involved. A resampling approach is usually implemented, e.g., leave-one-out or bootstrapping, that generally yields meaningful and realistic estimates of performance. The clinical value of a CADiagnosis method is usually assessed in two stages: First, computer performance is evaluated based on "ground truth" information provided by the experts and pathology using computer generated receiver operating characteristic (ROC) curves. Computer ROC curves are plots of true positive fraction (TPF) vs. false positive fraction (FPF) pairs that are generated by adjusting the algorithms' parameters. Publicly available software tools for ROC analysis may be used to fit the data and estimate performance parameters such as the area under the curve, A_Z, its standard error (SE), confidence intervals, and statistical significance. Following the laboratory evaluation, a true ROC experiment is usually performed that involves relatively large number of cases and human observers. The cost and time requirements of an observer ROC study are significant impediments in its implementation and such analysis is usually reserved for fully optimized techniques, namely for techniques that have been through rigorous computer ROC evaluation.