

Local spatial frequency image properties and radiological decisions

Mariusz W Pietrzyk^{a,b}, David J Manning^a, Tim Donovan^{a,c} and Alan Dix^b

^aSchool of Medical Image Sciences, University of Cumbria, Bowerham Road, Lancaster, LA1 3JD, ^bComputing Department, Lancaster University, South Drive, Lancaster LA1 4WA, ^cPsychology Department, Lancaster University, Fylde College, Lancaster LA1 4YF

Abstract. This study investigates the differences between local physical properties of regions selected by observers in terms of visual attention allocation and the decision-making process. The experience level is considered by involving the observers with different radiological background from expert to naïve. The spatial frequency analysis applies wavelet packet transforms. Dissimilarities are found in two comparisons both between the wavelet representations of the most dwelled TN vs. TP regions of interest, and also, between FP responses done by the subjects from different experience groups in terms of the analysis of variance at the $p < 0.05$ significance level, where the *post-hoc* tests are used. Firstly, the more experienced observers seem to distinguish TP better from all potential image feature candidates compared with less experienced observers, which may give insight into the unique experts' ability to avoid FP. Secondly, higher consistency of FP wavelet representations are found within a particular observer group than between two different groups. The farther two subjects are positioned from each other on the experience scale the greater significant differences are found in the physical description of overt errors on a normal region of interest. **Keywords:** *visual perception, radiological error, eye tracking, wavelets.*

1. Introduction

Human observers play an integral role in any medical imaging system, where radiological expertise requires a visual inspection of the output data. Because of this, the accuracy of imaging suffers from subjective limitations such as eye optics, photoelectrical translation capacity of the retina, image processing through neural pathways and higher cognitive processes, i.e. understanding developing through visual information analysis and interpretation involved comparison with previous examined cases (memory), expectations or satisfaction of search [1]. Observers faced with the task actively perform the visual searching and decision-making process depending on the perceived information. The interaction strongly depends either on the physical quality of image and the subject's individual features and attitude at the time of reading; e.g. experience, knowledge, intuition and visual judgment [2]. The radiological task is extremely difficult, involving searching for a target in very hectic environment, with little prior information about where to look (a background signal), and even less information about what to find (a target). State of the art imaging technology assists in reducing the image noise level, which makes easier to a better understanding of the background. Nevertheless, it seems unable to create a fully automated system design for visual searching, detection and recognition, which could take responsibility for providing radiological expertise and replace the clinician. Human observers still out-perform any artificial machine in the radiological task, although, subjects cooperating with the computer-aided detection (CAD) schemas serving as a second reader are reported to increase the accuracy of radiological performance in different tasks, e.g. lung nodule detection [3] or mammography [4]. Nonetheless, a better understanding of radiological error may bring additional tools in human error management [5].

A common radiological error occurs in chest radiographs inspection for lung cancer, where in approximately 30% of cases observers miss the lung nodule [6]. What is more, eye tracking studies, based on the retrospective experiments, report, that 60-70% of these errors are due to wrong judgment concerning the well localized abnormality – a decision-making stage, whereas only in 20% of missed nodules have not been gazed at all – a recognition stage [7]. The perceptual-cognitive skills to perceive the crucial information for right interpretation develops through years of practice, training, experience-gathering and knowledge validation. In consequence, expert performance differs from less experienced observers, by higher accuracy scores using fewer and shorter fixation points, covering smaller image area in the lung nodule detection from PA chest radiographs [8]. The background concept of visual perception as a decision-centred process [9] has been developed into a descriptive model of radiological image interpretation [10][11]. The model is based on the subject's fixation pattern committed during image investigation.

Neurophysiological and psychophysical findings show that multiscale transforms seem to appear in the visual cortex of mammals [12] so, visual perception depends upon, amongst other things, on frequency components of the modified contrast [13] and the orientation of the stimuli [14]. Hence, humans perceive the world through a number of visual channels such as colour and contrast, but some information is carried by independent spatial frequency ranges and orientations [15]. Recent studies have shown [16][17][18], that both local and global image-based elements have particular physical properties that are correlated with the performance and the level of experience of human

m.pietrzyk@lancaster.ac.uk

observers in medical imaging cancer detection. It seems likely, that some local image properties of the regions of interest have an influence on observer performance guiding their visual attention to certain locations, which results in their searching and sampling strategy during image reading [19][20]. Particularly, spatial frequency properties are hugely relevant to the human visual system and bring a comprehensive physical description of the set of selected sub-images. In our previous study [21], it has been found, that the most dwelled TN decision locations have greater similarity to nodule-contained ROIs for more experienced subjects in terms of spatial frequency properties. This study using eye tracking data collected during chest radiograph reading by observers with different radiological background attempts to answer the following questions: (1) is there a correlation between where a certain decision outcome had been made and the physical properties of that location? (2) are there differences between physical properties of those regions classified into the same decision outcome by observer with different experience?

2. Materials and Methods

40 observers are involved in the eye tracking experiments: ten experts with at least a few years experience in the chest radiographs examination; two levels of middle-experienced subjects (novices) recruited from 3rd and 2nd year of Radiography BSc at University of Cumbria with 28 and 12 weeks clinical experience and, 10 subjects not conversant with the task constitute the naïve group. Ten normal and 10 abnormal (with single or multiple nodule-contained cases) Posterior-Anterior chest radiographs (2⁸bits gray scale images) form the image data bank used in the experiment [22]. Twenty-three nodules ranged from 5 up to 30 mm in diameter. Subjects were informed that cases which they are going to investigate may contain single, multiple or none at all nodules. However, they were not made aware of the prevalence. Tobii x50 system and ClearView analysis software, with 50Hz sampling rate and 0.5 visual degrees space accuracy was used in the experiment, where images were displayed on a Medion 18.8" LCD monitor. This stand alone eye tracking device allows subject to keep a limited freedom of head movement, where any drift-effects are removed by implementation a binocular averaging.

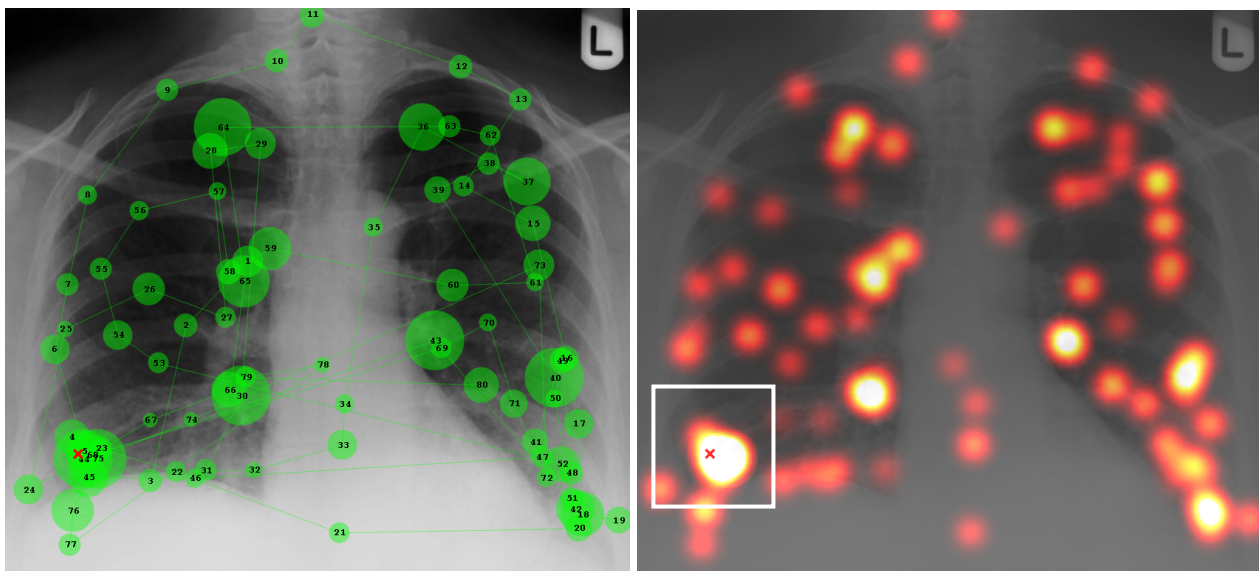


Figure 1. Eye tracking data: (*at left*) visual scan pathway pattern where time ordered fixations (green circles with radius indicating dwell time) are linked from first to the last gazed location. Cross represents the mouse click location, which marks the overt decision location (in this case it is a TP); (*at right*) hot spot where colour scale indicates the cumulative dwell time: white = above 1 sec, red = above 330ms. The selected ROI frames the nodule-contained area.

It is assumed, that eye movement is synchronously shifted with spatial-visual attention [23], and therefore saccades and fixations gain insight into the overt localization of visual attention. However, only that information which was perceived during fixation is transmitted to and processed in the visual cortex [24]. Because of this and for the purpose of this study, a filtering procedure was implemented to exclude saccades from data and a time and space clustering algorithm was used to find the fixation locations. The dynamics of the changes to the visual axis during visual search is described by the acceleration and velocity; and the fixation threshold is set up to 8000°/sec² and 30°/sec respectively. The spatial relative distance threshold fixation from saccades equals to 50 pixels. This equates to a defined subject's fixation upon a particular local region of interest as a gaze area limited by a foveal angle of 2.5 degrees, dwelled for at least 100ms. The time ordered fixation pattern depicts the visual scan path of subject's searching performance (Figure 1 – *at left*). Next, the cumulative dwell time map (Figure 1 – *at right*) with local maxima and their relative distance from nodule and mouse click marks were used to localize covert and overt observer's responses according to the applied search model [25]. The ignored regions were defined by 330ms

threshold. The overt response indicated by mouse clicks may be classified as true-positive (TP) or false-positive (FP) depending on the nodule present or absent respectively in the ROI. The missed nodules are scored as false-negatives (FN). The true-negative TN occurs where a normal ROI is the dwelled over 330ms left unmarked. Nonetheless, only the most dwelled TN regions per each case are analysed in this study.

The local background around a particular location seems be crucial in the visual searching and decision-making process [26]. The locality of these regions of interest (ROI) is defied by the foveal visual angle (which is approximately a 100 pixels box viewed from 60-70cm). In this way, image processing techniques were implemented to analyse the physical properties of local background selected locations – at decision-point sites. In particular, the visual information, as signal energy, carried by spatial frequencies $\omega(\omega_x, \omega_y)$ in a certain orientations $\theta = \tan^{-1}(\omega_x/\omega_y)$ are the main interest of this research. The Wavelet Packet Transformation (WPT) algorithm [27] has been applied to spatial frequency analysis of selected ROIs up to the 3rd level of decomposition frames by Daubechies functions. This band-pass filtering iteration algorithm is designed to perform hierarchical transformations, where obtained from the previous level ($j-1$) sub-images are transforming again as an input signals at level j and forward to further analysis until the last level of decomposition (Figure 2). Each sub-image $f_j(x,y)$ from a particular level j represents different information called a spatial frequency band (SFB), and so, the approximations a_j were obtained as a low spatial frequency representation of the input ROI: $\forall_{j \in \mathbb{Z}} a_j = \langle a_{j-1}, \phi_j(x)\phi_j(y) \rangle$, whereas high spatial frequency components are contained by sub-images $f_{j-1}(x,y)$ represented details $d_{i=1:3}^j$ in a particular orientations (horizontal $_{|i=1}$, vertical $_{|i=2}$, diagonal $_{|i=3}$): $\forall_{j \in \mathbb{Z}} \sum_i d_i^j = \langle f(x,y), \psi_j(x)\phi_j(y) \rangle + \langle f(x,y), \phi_j(x)\psi_j(y) \rangle + \langle f(x,y), \psi_j(x)\psi_j(y) \rangle$. As a result, 84 bands were obtained for each ROI. Each band has a physical meaning and contains a certain amount of information, which has been quantified according to the logarithm of energy within that sub-image: $E = \log(\sum_{n,m} |f(x_n, y_m)|^2)$. That value per band was used in further statistical analysis. The amount of texture similarities was quantified according to the ANOVA of energy within SFB considering the type of decision and the author level of experience at the statistical significance $p < 0.05$ with Scheffe post-hoc test [17]. The number of SFB with significance difference between particular factors is interpreted as a dissimilarities indicator. The more bands are different the less similar are two compared representation.

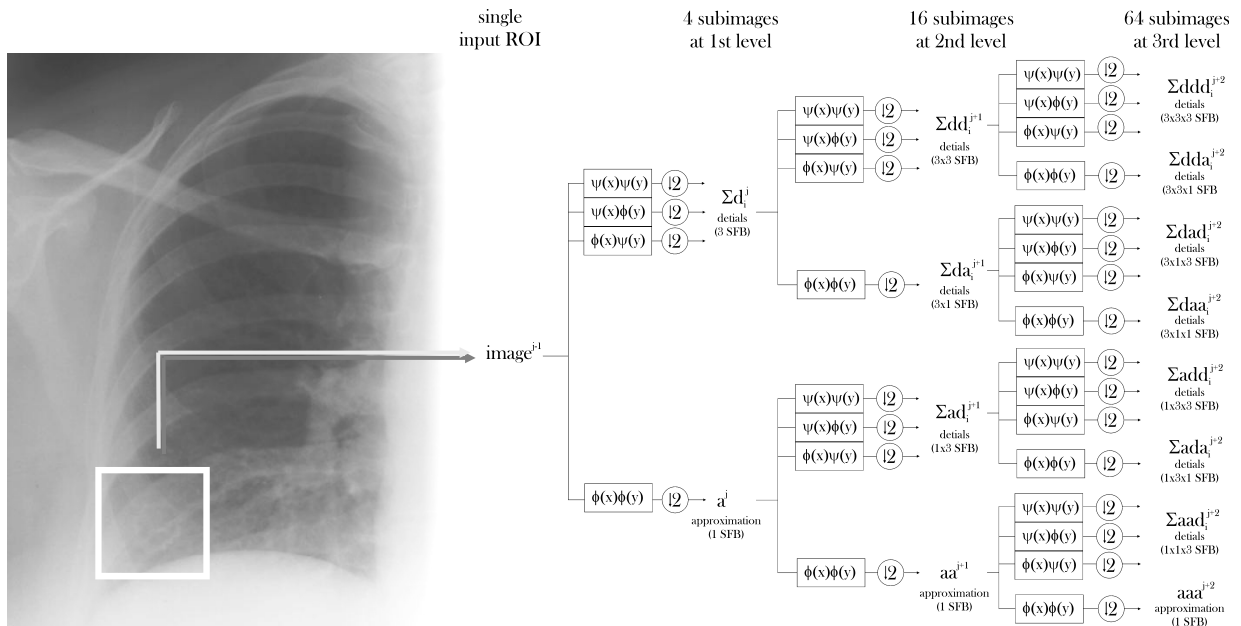


Figure 2. Schema of the Spatial Frequency Analysis of selected ROI. Wavelet Packet Transforms up to 3rd level of decomposition were applied to frame the input 2D signal into 4+16+64 spatial frequency bands, each characterized by different range in spatial frequency at various orientations. Approximations of the signal were obtained from 2D low-pass filtering in both orientations (rows(y) and columns(x)): $F(x)F(y)$, whereas details obtained from high-pass Y filtering: $F(x)Y(y)$, $Y(x)F(y)$, $Y(x)Y(y)$, represent high frequency components at particular scale and orientation. $\downarrow 2$ indicates the down-sampling step in transformation algorithm.

3. Results and discussion

Radiological performance accuracy is characterised by the number of outcomes, where the subject's decision are compare with ground truth. The eye tracking data on retrospection allows allocating not only overt decisions (TP, FP) but also covert decisions: FN and TN. The first research question is addressed to the (dis)similarities in spatial frequency description according to the wavelet representation between local regions of interest related with different

radiological decisions: the most attractive TN and correctly reported pulmonary nodules (TP). The number of SFBs which are significant different at $p < 0.05$ level of confidence according to ANOVA was used to compare Spatial Frequency Analysis results, which differ between observer groups (Figure 3). The analysis is an attempt to find out, which spatial frequency properties have been tuned in order to correctly distinguish the suspicious object as a truly normal or abnormal region in terms of lung nodule recognition. The presented trend in the statistical nature of particular bands has been observed across radiological experience, where more SFBs with significant differences occurs for more experienced subjects up to the expert level.

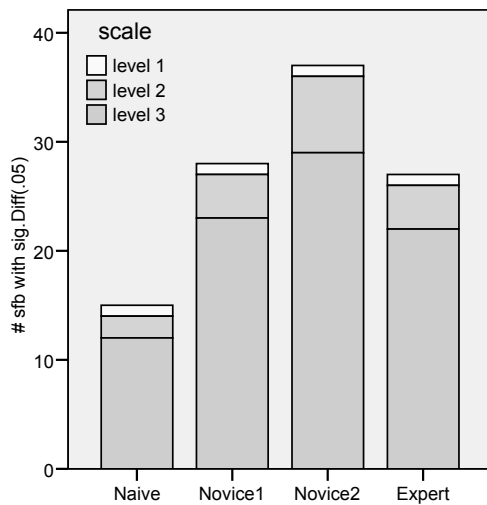


Figure 3. The statistical pair comparison results of the most dwelled TN and TP wavelet representation.

Paired subjects' experience groups		Number of SFB with significant difference
Experts	Novice2	3
Experts	Novice1	19
Experts	Naive	26
Novice2	Novice1	0
Novice2	Naive	7
Novice1	Naive	0

Table 1. The statistical comparison of FP wavelet representations up to 3rd level marked by subjects form different radiological experience according to the ANOVA with Scheffe post hoc test. The number of SFB indicates the texture differences at the significance level $p < 0.05$. Overt errors differ between the subject's experience group, and that difference is more obvious, if the distance at the experience scale is bigger.

The second question refers to the normal regions, where subjects misinterpret suspicious (but normal) objects as lung nodules based on their visual judgment of physical image features. These false-positive decisions (FP) captured a lot of visual effort measured by relative long cumulative dwell time spent on them. The FP regions of interest were compared between the different observers. The results indicate that the FP sites, unified in terms of spatial frequency properties, simultaneously correlate with radiological experience. It seems that subjects from the same radiological background tend to misinterpret the visual information in a more similar way compared with the people with different experience (Table 1). In other words, the farther the separation distance is between two subjects on the experience scale, the more obvious physical differences will be observed between their FP types. The more closely two subjects are in the experience background, the less differences are observed in spatial frequency properties between their FP types.

4. Conclusion

The perceptual approach to radiological error in lung nodule detection from PA chest radiographs was implemented in this study through eye tracking methodology. Wavelet decomposition is used to analyse the spatial frequency properties of selected region of interest, which were classified relative to the ground truth into decision category: TP, FP, TN or FN. The statistical comparison (ANOVA, $p < 0.05$) of these decision-site sub-images objectives, i.e. spatial frequency properties, outcome type, the amount of cumulative dwell time spent on it and the experience level of subject brings insight into the mechanisms underlying medical image perception.

The experts seem to perform better because of their ability to distinguish TP from all the most attractive TN, which results in higher accuracy by avoiding the FP responses. Spatial frequency analysis gained insight into that phenomenon, showing the set of features based on those two types of regions which have been perceived differently. The physical description of local sub-images and the types of the decision made upon these regions across the experience level seems to be correlated. What is more, the visual attention distribution seems to be correlated as well in a very similar way, which may bring additional insight into the definition of a radiological saliency function that depends on the type of the task, individual subject's features and display properties.

The similarities in FP between more experienced observers may suggest that unification in terms of spatial frequency properties originate together with experience gathering and perceptual-cognitive skill development in being able to detect and recognise a particular target. The correlations between levels of experience in the radiological task are a work still in progress.

It is well known and expected that experts perform better than non-experts. Medical image perception studies have developed the metrics to measure and describe subjects' accuracy in visual searching, detection and recognition tasks. However, the reasons for better performance in the radiological tasks are still being hypothesised. Recent studies provide support for one of the potential explanations of the quality in expert performance through the visual channels theory. According to that hypothesis, experts may develop and use specific neural connections – spatial frequency channels tuned to specific objects detection – during visual searching in a radiological task. These visual pathways are used like convolution filters tuned to relevant information. This hypothesis is supported by findings in the speed of expert behaviour, which may suggest that the early stage of visual information processing is too fast to involve cognition and may be based on the prior selection from the neurological network.

References

- 1 K.S. Berbaum, E. A. Franken, D. D. Dorfman et. all "Satisfaction of search in diagnostic radiology", *Investigative Radiology* **25**(2), pp. 133-140, 1990.
- 2 D. Manning, A. Gale & E. Krupinski "Perception research in medical imaging", *The British Journal of Radiology* **78**, pp. 683-685, 2005.
- 3 K. Doi "Computer-aided diagnosis in medical imaging: historical review, current status and future potential", *Computerized Medical Imaging and Graphics* **31**, pp. 198-211, 2007.
- 4 S.H. Taplin, C. M. Rutter & C. D. Lehman "Testing the effect of computer-assisted detection on interpretive performance in screening mammography", *American Journal of Roentgenology* **187**, pp. 1475-1482, 2006.
- 5 R. Fitzgerald "Error in radiology", *Clinical Radiology* **56**, pp. 938-946, 2001.
- 6 D.P. Carmody, C. F. Nodine & H. L. Kundel "Finding lung nodules with and without comparative visual scanning", *Perception and psychophysics* **29**, pp 594, 1981.
- 7 D.J. Manning, S. C. Ethell & T. Donovan "Categories of observer error from eye-tracking and AFROC data", In *SPIE Medical Imaging: Image Perception, Observer Performance, and Technology*, pp. 90-99, SPIE Library, 2004.
- 8 D.J. Manning, S. C. Ethell, T. Donovan et. all "How do radiologists do it? the influence of experience and training on searching for chest nodules", *Radiography* **12**, pp 134, 2006.
- 9 I. Rock. *The logic of perception*. The MIT Press, Cambridge MA, 1983.
- 10 C.F. Nodine & H. L. Kundel "Using eye movements to study visual search and to improve tumour detection", *Radiographics : a review publication of the Radiological Society of North America, Inc* **7**, pp 1241, 1987.
- 11 H.L. Kundel, C. F. Nodine & E. A. Krupinski "Computer-displayed eye position as a visual aid to pulmonary nodule interpretation", *Investigative Radiology* **25**, pp 890, 1990.
- 12 D. Hubel. *Eye, Brain, and Vision*. <http://hubel.med.harvard.edu>, 1995.
- 13 F.W. Cambell & D. Green "Optical and retina factors affecting visual resolution", *Journal of Physiology (London)* **181**, pp. 576-593, 1965.
- 14 F.W. Cambell & J. Kulikowski "Orientation selectivity if the human visual system", *Journal of Physiology (London)* **197**, pp. 437-441, 1966.
- 15 F.W. Cambell & J. G. Robson "Application of fourier analysis to the visibility of gratings", *Journal of Physiology (London)* **197**, pp. 551-566, 1967.
- 16 C. Mello-Thoms, S. M. Dunn, C. F. Nodine et. all "An analysis of perceptual errors in reading mammograms using quasi-local spatial frequency spectra", *Journal of Digital Imaging* **14**, pp. 117-123, 2001.
- 17 C. Mello-Thoms, S. M. Dunn, C. F. Nodine et. all "The perception of breast cancer: what differentiates missed from reported cancers in mammography?", *Academic Radiology* **9**, pp. 1004-1012, 2002.
- 18 C. Mello-Thoms & B. Chapman "A preliminary report on the role of spatial frequency analysis in the perception of breast cancers missed at mammography screening", *Academic Radiology* **11**, pp. 894-906, 2004.
- 19 C. Mello-Thoms "The 'holistic grail': possible implications of an initial mistake in the reading of digital mammograms", In *SPIE Medical Imaging: Image Perception, Observer Performance, and Technology Assessment*, pp 20, SPIE Library, 2009.
- 20 H.L. Kundel & C. F. Nodine "Modelling visual search during mammogram viewing", In *SPIE Medical Imaging: Image Perception, Observer Performance, and Technology*, pp. 110-115, SPIE - The International Society for Optical Engineering, 2004.
- 21 M. Pietrzyk, D. Manning, A. Dix & T. Donovan "Spatial frequency characteristics at image decision-point locations for observers with different radiological backgrounds in lung nodule detection", In *SPIE Medical Imaging: Image Perception, Observer Performance, and Technology Assessment*, pp 17, SPIE Library, 2009.
- 22 S.B. Ethell. *Disease prevalence: an influential factor in radiological performance*. School of Medical Imaging Sciences, Lancaster, 2003.
- 23 J.E. Hoffman. Visual attention and eye movemnets. In *Attention*. Pashier (Ed.). 1998. .
- 24 A. Duchowski. *Eye Tracking Methodology: Theory and Practice*. Springer-Verlag, London, 2007.
- 25 D.P. Chakraborty "A search model and figure of merit for observer data acquired according to the free-response paradigm", *Physics in Medicine and Biology* **51**, pp. 3449-3462, 2006.
- 26 C. Mello-Thoms "The problem of image interpretation in mammography: effects of lesion conspicuity on the visual search strategy of radiologists", *The British Journal of Radiology* **000**, pp. 1-6, 2006.
- 27 S. Mallat "Wavelets for a vision", In *IEEE*, pp. 604-614, 1996.