A Methodology for Detecting Vessels in X-Ray Mammogram Images

Abstract: Identification of the blood vessels and milk ducts within an X-ray mammogram image allows intelligent analysis of calcification clusters and mass borders detected by other means. These vessel features are however generally buried within noise with variance of the order of the vessel signal itself. We present a scale-matched feature detector designed to extract nominally 1-dimensional ridges from an image with a signal-to-noise ratio approximating unity. The algorithm uses first and second difference gradient measures of the original non-smoothed images to drive a search exploiting the priori knowledge of the expected vessel (ridge) features. The algorithm is illustrated with an example.

1 INTRODUCTION

Screening of all women for breast cancer in the high risk groups of 50–65 years requires annual assessment of three million X-ray films in the UK. Computerised analysis of a mammogram film can potentially improve the assessment of the film by giving the diagnostic radiologist additional information upon which to base a decision.

In assessment of a mammogram film, it is desirable to be able to identify the vessels (both blood vessels and milk ducts, hereafter referred to simply as vessels) within a mammogram image. Some of the most important mammographic indicators of breast cancer are asked to their relative location and interaction with the vessels of the breast.

For instance, the degree of importance of a cluster of calcifications (arguably the most important mammographic indicator of cancer) is related to their localisation within a single vessel, or their spread over a number of adjacent vessels [Caseldine et al., 1988]. The actual detection of calcifications has been the subject of much fort and this issue is not addressed in this paper.

Another important mammographic indicator of developing tumours is a spiculated mass. The developing tumour requires blood for survival and growth and accordingly develops a series of blood vessels which "feed" the tumour. The vessels approach the tumour from all directions, resulting in a mammographic signature of central mass surrounded by a series of radial vessels, rather like the spokes of a bicycle wheel. In determining a potential mass is a spiculated mass it is important to simply assess the "roughness" of the mass border, but also to determine if the vessels approaching the mass terminate in the mass, or pass completely through it (2-D projected) mass [Caseldine et al., 1988, 1992]. Knowledge of the vessel structure in an image allows the assessment of vessel and mass interaction and the degree to which the vessels "feed" the masses.

Thus there is a range of diagnostic information than can be extracted from a mammogram image given knowledge of the corresponding vessel network. This paper briefly describes a system developed to extract the vessel network from a mammogram image. Section 2 describes the characteristics of the vessels and the images in which they are found, developing the constraints that the feature detector must accommodate. Section 3 looks at the traditional computer vision approaches to detecting features, identifying the shortcomings in relation to the present task. Section 4 initially states the vessel detector algorithm and then briefly reviews its components, while section 5 gives results of application of the algorithm to mammogram images.

2 TASK IDENTIFICATION

When identifying the attributes of a successful vessel detector it is instructive to initially list the features of a vessel as presented in a typical digital image.

Imaged vessels present mammographically as regions of marginally lower film density (due to the increased X-ray beam attenuation due to the marginally higher material density along the X-ray path which traverses a vessel), and in the digitised image are thus of higher absolute grey-scale value. Thus we are interested in looking for nominally lighter regions against a darker background. A cursory glance at a mammogram image, for example that of figure 1(a), reveals that vessels are nominally 1-dimensional structures with low local curvature. By inspection, it is revealed that the absolute grey level difference between those pixels deemed by eye to belong to a vessel-like feature and those adjacent pixels that do not (the background) is often as low as ~1.6% of the available quantised intensity range (~4 intensity levels in an 8-bit quantisation). In Appendix A we show that the assumed gaussian image noise has variance $\sigma^2 \approx 1.2$–2% (3–5 grey levels at 8-bit resolution) of the intensity range, giving a signal-to-noise ratio (SNR) approximating unity in the worst case. There exists however, sufficient information to distinguish vessel pixels from background, since one can do it by eye. Clearly this information lies in the fact that the vessels are connected, and it is this structural information that separates the wheat from the chaff.

Unlike many typical industrial computer vision applications, the adjacent features (vessels) of these images may lie along side one another at a spacing of only a few pixels (given the digitisation resolution). Thus, in
Physiologically the majority of vessels range from capillaries up to \( \sim 1\text{-}2 \text{mm} \) diameters. The spatial extent of an imaged vessel in a direction perpendicular to its local axis is strictly dependent upon the digitisation resolution. We currently use mammogram images with a film resolution of \( 300 \mu\text{m} \) square per pixel, meaning that most vessels will appear at widths up to \( \sim 6 \) pixels.

In summary we require a system that can detect signals in unsmoothed images with \( \text{SNR} \approx 1 \). The vessels present typically as nominally 1-dimensional structures at up to \( \sim 6 \) pixels in width, with increased pixel intensity relative to the background or surrounding area that can be as low as 1.6% of the total intensity domain.

3 TRADITIONAL COMPUTER VISION TECHNIQUES

Detection of edges and ridges in images has been the subject of widespread investigation in the computer vision community. All algorithms, at their simplest level, build a model of the feature to extract and look for existences of the model in the image. A vessel can be modelled (i) as a pair of back-to-back step transitions, or (ii) directly as an indivisible structure with a cross-section of a top-hat or a hump.

The first case is initially attractive because it suggests that conventional step/edge detectors followed by ridge finding using a logical combination of step pairs would be successful. In the case of mammogram images, many of the standard methods, such as Canny’s gradient-based edge detector [1986], fail due to the low SNR and their need to pre-smooth the images, generally as a noise suppression technique. As an example of this, figure 2(b) shows the results of Canny’s edge detector applied to the image of figure 1(a). Morphological edge detectors fare no better, a result that is tied directly to the structuring element, which cannot distinguish between noise and feature. Additional issues arise in the selection of the structuring element itself — is a spherical, ellipsoidal or other element best suited to the expected edge profile?

The second case, of directly modelling a ridge has been the subject of little attention in the computer vision literature since attempts to generalise the model of a ridge in order to develop an algorithm leave it oversimplified and incapable of capturing the peculiarities of most practical applications. Consequently most successful implementations are application specific. An example of a general ridge finder is the ridge/ravine finder of Haralick [1992] based on the bicubic facet model. The results of applying this method to the image of figure 1(a) are shown in figure 2(a). The poor response is mostly attributable to the low SNR. Clearly these general methods are inappropriate for the present task.

Fleck [1988] describes an edge finder system, named Phantom, based on topological interpretation of second difference measures of an unsmoothed image. In the broadest generalisation possible, the Phantom edge finding process can be summarised as in figure 3.

- A. Determine the second difference responses in a number of directions.
- B. At each pixel, combine the information from the second differences calculated in the various directions, segmenting the image into pixels with significant positive, significant negative, or insignificant response.
- C. Correct for errors due to noise and other artifacts.
- D. Hypothesise boundary locations that explain the observed second difference response patterns.

The secret to Phantom’s success lies in the details...
the algorithms to combine the directional information and suppress the errors. By combining both shape and amplitude information into a single summation over the maximal star-convex neighbourhood surrounding each pixel, the algorithm can distinguish real responses from those due to noise. It is within this framework of noisy feature detection and subsequent noise suppression based on careful interpretation of the data that the present work is grounded.

4 THE MAMMOGRAM VESSEL DETECTOR

The first instance the framework of the algorithm is applied and this is followed by a brief discussion of its composition.

1 The Ridge/Vessel Detector Algorithm

At Δν' and Δν'' be the first and second difference filters respectively, where α = {N, NE, E, SE} representing the north, north-east, east and south-east directions respectively (see figure 4). Let the convolution of these filters with the image I be represented as Iν' = Δν'* I, and Iν'' = Δν''* I. Let I(i,j), Iν'(i,j) and Iν''(i,j) be the image value at the pixel with coordinates (i,j) of the respective image. Using this notation, the algorithm is us:

Form the eight difference images Iν' and Iν''.

Find the gradient of the image surface, ∇I, using the first difference images Iν'. Since the image I is not pre-smoothed before calculating Iν', the spatial extent of the Δν' filters is so small (2 pixels), there are numerous local errors in Iν' that deem the simple summation at each pixel of the appropriate Iν' components of ∇I susceptible to unknown corruption. It is therefore necessary to consider the available evidence for an image gradient at each pixel.

At each pixel (i,j) in I, take the eight Iν' values that involve the pixel in their calculation and form the vectors

\[ H = \begin{bmatrix} I_N(i,j) \oplus I_N(i-1,j) \\ I_E(i,j) \oplus I_E(i,j) \\ I_S(i,j) \oplus I_S(i,j) \\ I_SE(i,j) \oplus I_SE(i,j) \end{bmatrix} \]

\[ D = \begin{bmatrix} I_{NE}(i,j) \oplus I_{NE}(i-1,j-1) \\ I_{SE}(i,j) \oplus I_{SE}(i,j-1) \end{bmatrix} \]

where the coordinate system of D is that of H rotated through \( \frac{\pi}{2} \) radians, the \( \oplus \) operator returns the maximal response of the two arguments such that if they are of the same sign then return the algebraic sum of the arguments, otherwise return the argument with the maximum absolute value.

If a consistent gradient vector occurs over the neighbourhood of the pixel, then the vectors H and D should be roughly aligned (with D expressed in the coordinate system of H). Thus if the included angle is below some threshold, then there is consistent evidence across the neighbourhood for the existence of the image gradient, and it is thus set:

\[ \nabla I(i,j) = \begin{cases} H + D & \text{if } \arg(H) - \arg(D) < \theta_{\Delta'}, \\ 0 & \text{otherwise}. \end{cases} \]

3. Select all pixels with a significant Δ'' response as feature candidates by forming the set N as:

\[ N = \{ (i,j) : \max(I''_N(i,j), I''_{NE}(i,j), I''_E(i,j), I''_{SE}(i,j)) > T_{\Delta''} \} \]

4. Scan the set N, omitting any pixels that form 8-connected regions of fewer than a given number of pixels.

5. Take the remaining 8-connected regions of set N and thin them down to a simply connected skeleton that is approximately on the medial-axis of the regions. Form the set S of pixels lying on these region skeletons.

6. Track along the elements of set S linking together into bones b those elements that lie in the same direction. Upon encountering a branch in the skeleton, initiate another bone. Continue classifying the elements of set S until no further unclassified elements remain. Reject those bones of insufficient length, forming the set B of all retained bones b.

7. Follow the axis/skeleton of each bone b ∈ B. Searching in a direction perpendicular to the local axis (heading), search the set N for 8-connected pixels

\[ \theta_{\Delta'} \approx \frac{\pi}{2} \] to allow for the variation due to the noise. Note that this threshold does not effect the value of the image gradient calculated at (i,j), rather it defines the point at which the available evidence for an image gradient is, or is not believed.

\[ T_{\Delta''} \approx 70-80\% \] of the standard deviation \( \sigma \).

\[ \Delta' \]

\[ \Delta'' \]

\[ 1 \]

\[ \Delta_N \]

\[ \Delta_{NE} \]

\[ \Delta_E \]

\[ \Delta_{SE} \]

\[ 1 \]

\[ \sqrt{2} - 2 \]

\[ 0 \]

\[ 1 \]

\[ \Delta''_N \]

\[ \Delta''_{NE} \]

\[ \Delta''_E \]

\[ \Delta''_{SE} \]

\[ 1 - \sqrt{2} \]

\[ 0 \]

\[ 0 \]

\[ 1 \]

\[ \Delta''_N \]

\[ \Delta''_{NE} \]

\[ \Delta''_E \]

\[ \Delta''_{SE} \]

\[ \sqrt{2} - 2 \]

\[ 0 \]

\[ 0 \]

\[ 1 \]

\[ \Delta''_N \]

\[ \Delta''_{NE} \]

\[ \Delta''_E \]

\[ \Delta''_{SE} \]

\[ 0 \]

\[ 0 \]

\[ 0 \]

\[ 1 - \sqrt{2} \]

\[ 0 \]

\[ 0 \]

\[ 0 \]

\[ 1 \]

The difference operators Δ' and Δ''. Note that [Δν''] = [Δν']T and [Δν''SE] is formed by mirror folding [Δν''NE] about its central column. The results of the ' convolution are assigned to the pixel at the top left corner of the mask, whilst the Δ'' result is assigned to the central pixel of the mask.
that surround the skeleton pixels of bone $b$. Observing the image gradients calculated at step 2 retain in the vessel $v$ those pixels that do not contradict the notion of a vessel with the local heading of $b$. Initiate a new vessel $v$ for each bone of set $B$ encountered, and continue until $B$ is exhausted.

8. Declare the set of all vessels $v$ to be the “vessels” of the image $I$.

4.2 The Algorithm Briefly Reviewed

Since the vessels are of higher intensity and are typically up to $\sim 6$ pixels in widths, the second difference operators $\Delta'^2$ of figure 4 (which are actually scaled, inverted second difference operators) with spatial extent of 5 pixels were chosen to match the spatial characteristics of the expected vessel profile, rather than for the general computer vision reasons of image interlacing (which is not present in these images due to the digitisation process). Step 3 (which equates to step B of figure 3) collects only those pixels that are potential/candidate vessel pixels, whilst steps 4–8 introduce knowledge of the expected vessel profile to rule out artifacts (steps C and D of figure 3). It is in these steps that the direct modelling of the vessel profile is introduced into the system.

The exact details of the various implementations (e.g. the thinning algorithm of step 5) are superfluous to the methodology presented in section 4.1 and are omitted for simplicity.

5 Results

Although it is somewhat difficult to present quantitative results, for completeness the result of running the algorithm over the image of figure 1(a) is given in figure 1(b). Similar results are obtained for images digitised across a variety of platforms. The strength of this algorithm is that the feature description allows many other tasks to be undertaken. For example, “removing” the vessels gives an image freed from the textural clutter of the vessels, allowing improved detection and analysis of masses, and improved tissue classification. Combining calcification data with the vessel description (mainly the local vessel direction) improves differentiation between arterial/venous calcifications and ductal calcifications.

6 Conclusion

We present the framework of an implemented algorithm for extracting a description of the vessel features from X-ray mammogram images. The algorithm is scale-space matched to the expected cross-section of the vessels, and whilst the algorithm has not been implemented at varying scales, there is no reason why this could not be done. The algorithm can successfully differentiate spatially adjacent features, and resolve features with a SNR approximating unity.

Current work is directed towards developing clinically significant tools using the vessel description to analyse calcification clusters, masses (in particular spiculated masses) and bilateral asymmetry using both the vessel networks themselves and the texturally less cluttered “removed-vessel” images for comparison.

A Feature Signal-to-Noise

Digital X-ray mammogram images are typically attained by digitising X-ray film mammograms (although the technology exists for direct digital imaging, this is not likely to make a significant clinical impact for some time). The digitisation process itself introduces white gaussian noise, predominantly shot noise within the imaging device, in addition to any noise introduced during the X-ray exposure/developing of the mammogram film.

In order to determine the average noise level of an image, consider the facet-based peak noise model of Haralick [Haralick and Shapiro, 1992]. Following Haralick, fit a sloped facet approximation to the image surface, and determine the squared differences between the fitted surface and the image as:

$$\epsilon_{(i,j)}^2 = \sum_{(r,c) \in N} \left[\alpha r + \beta c + \gamma - I(r,c)\right]^2$$

where $N$ is the 8-connected region over $(r,c)$ surrounding the central pixel $(i,j)$ over which the facet-model is fitted to the data, and $I$ is the raw image data. A least-squares minimisation of $\epsilon^2$ gives the model parameters $\alpha$, $\beta$ and $\gamma$ at each $(i,j)$. Each neighbourhood’s normalised squared residual error $\epsilon^2/(\sum_r \sum_c (1-3))$ can constitute an unbiased estimator for the variance $\sigma^2$ of the noise. When averaged over all the pixels in an image this is a stable estimator of $\sigma^2$. In the present example, this estimator is further stabilised by only calculating $\epsilon^2$ over neighbourhoods consisting entirely of background pixels, since the sloped-facet model in addition to assuming a gaussian noise process also assumes that the underlying real image surface is a piecewise linear surface over the pixel neighbourhood. This assumption finds more support in background regions.

Implementation of this method over many images of varying conditions and digitisation processes shows that the background noise is of zero mean and variance $\sigma^2 \approx 3$–5 intensity levels.

Acknowledgements

NC acknowledges the support of the Rhodes Trust.

References