Detection of implanted markers in radiographic image sequences

M. Egmont-Petersen, T. Arts

Department of Biophysics
Maastricht University
P.O.B. 616, NL-6226 EL Maastricht
The Netherlands
michael.egmont@bf.unimaas.nl

Abstract
In this paper, a nonlinear filter is developed for detection of radio opaque markers. It is shown that linear filters are only optimal under specific conditions. As these are not necessarily met in our case, it is suggested to use a nonlinear neural network for detection of markers. An experiment with simulated markers in which the neural network and the linear filter are compared, indicate that the neural network outperforms the linear filter. Another neural network is developed to predict the centre of the marker. The deviation of its predictions are analyzed in a second experiment.

Keywords: Object detection, template matching, neural networks, minimal error-rate classifier

1 Introduction

In investigating tissue deformation, radio opaque markers may be implanted in various animal tissues. This technique is used to study cardiac motion and the much slower growth of the cardiac muscle. The markers (usually metal spheres) are implanted in the cardiac wall of an animal. Data are collected by exposing the animal to X-rays in sequences of 5-10 seconds. The X-rays are projected onto an image amplifier which is scanned by a CCD-camera with 50Hz. The markers appear as dark spots because they absorb more X-rays than the surrounding tissue (see figure 1).

Automatic detection of the markers is problematic. It is impeded by variations among the markers in signal strength and shape of the shadow. When a marker is projected over a background of ribs and other tissues, its signal strength may decrease five fold.

Existing filter methods such as convolution and template matching can be used to detect markers. Such linear filters perform a convolution with a kernel that resembles the type of object one wants to identify. Application of linear filters does not always lead to an optimal detection result [4]. We indicate under which conditions linear filters are optimal. As these conditions are not necessarily met in our situation, we suggest to use a nonlinear neural network for detection. Provided with the intensities in a window that is slided across an image, the neural network acts as a nonlinear filter. Theoretical results from statistical pattern recognition indicate that a neural network will under certain conditions constitute the optimal filter.

Convolution filters including neural networks result in a detection of a marker with an accuracy of a half of a pixel. To obtain a more precise estimate of the position, we suggest to use a neural network to estimate the coordinates of the marker within the window. This boils down to a nonlinear regression task.

The neural networks for detection and nonlinear regression are trained using images with simulated markers to which Gaussian noise with a zero mean is added.

2 Background

An image can be seen as a 2–dimensional signal I sampled in a grid. Detection of an object can be obtained by filtering this signal.

2.1 Filtering as detection method

A vast variety of filter techniques exist. For detection of rotation invariant objects in an image, linear convolution and template matching are standard techniques. A convolution of an image I with a kernel C, which contains a bandlimited signal in the frequency spectrum, reduces signals with other frequencies, $Y = C * I$. For the coordinate pair (a, b) the convolution can be written as

$$Y(a,b) = \int_{a'min}^{a'max} \int_{b'min}^{b'max} C(a',b') \cdot W(a-a',b-b') \;\; da' \, db' \tag{1}$$

The kernel C resembles the intensity pattern of the type of object searched for, i.e. a marker. W is a window with r rows and l columns that is slided across the image. The resulting convolution value $Y(a,b) \in \Re$ is assigned to the pixel with the coordinate (a, b).

Template matching entails first a convolution with a kernel C and subsequently a normalization step. It results in the correlation coefficient between the window and the kernel. The convolution value is divided by the product of the signal strengths present in the window and the convolution kernel

$$Y(a,b) = \frac{\int_{a'min}^{a'max} \int_{b'min}^{b'max} C(a',b') \cdot W(a-a',b-b') \;\; da' \, db'}{\parallel C \parallel \, \parallel W(a,b) \parallel} \tag{2}$$

with $\parallel C \parallel$ defined as $(\Sigma_{a'}\Sigma_{b'} \; C(a',b')^2)^{0,5}$, $Y(a,b) \in (-1,1)$. Both convolution and template matching are linear filters.

A disadvantage of template matching in our application is that the amplitude of the filtered signal is disregarded. This causes problems because the texture pertaining to background (cardiac tissue) sometimes resembles a marker but the signal amplitude of the texture is an order of magnitude smaller than that of a marker. In general, linear filtering is known to be error prone and, as we will show, only optimal under specific conditions.

2.2 An image as stochastic signal

The pixels in an image I can also be seen as successive realizations of a random variable x, the intensity, that can take values in the range (α, β), with α the minimal and β the maximal intensity that can possibly be observed. The convolution (1) can be seen as a sample function which to each coordinate pair (a, b) results in a sample $W(a, b)$ of the image I. Filtering boils down to a statistical classification task in which each sample

$W(a,b)$ is classified to one of the two classes Marker or Noise. Linear filters such as (1) and (2) implement linear discriminant functions.

2.3 Restrictions on linear discriminant functions

Within statistics, it is well-known that linear discriminant functions will only in specific situations result in optimal discrimination between two classes. Let the two types of signals to be discriminated be M (Marker) and N (Noise) with the prior probabilities $\mathrm{P}(M)$ and $\mathrm{P}(N)$. Denote with $\mathrm{p}(W(a,b)|M)$ and $\mathrm{p}(W(a,b)|N)$ the probability density functions of the two signals observed through $W(a,b)$. A minimal error-rate classifier – the optimal classifier – assigns always the most likely class label to $W(a,b)$. The region of the minimal error-rate classifier in which the posterior probability is maximal that the observed sample $W(a,b)$ belongs to the class Marker, is given by

$$R_M = \{W(a,b) \in \Re^{(r \times l)} \mid \mathrm{P}(M)\mathrm{p}(W(a,b)|M) > \mathrm{P}(N)\mathrm{p}(W(a,b)|N)\} \qquad (3)$$

and the corresponding region R_N in which the probability that $W(a,b)$ is not a marker is $\Re^{(r \times l)} \setminus R_M$. Linear discriminant functions are only optimal when the regions R_M and R_N are separated by a plane. This is the case when the two distributions $\mathrm{p}(W(a,b)|M)$ and $\mathrm{p}(W(a,b)|N)$ are Gaussian and have identical covariance matrices Σ, $\mathrm{p}(W(a,b)|M) = \mathrm{g}(W(a,b)|\mu_M, \Sigma_M)$ and $\mathrm{p}(W(a,b)|N)= \mathrm{g}(W(a,b)|\mu_N, \Sigma_N)$, with $\Sigma_M = \Sigma_N$ [3]. In other situations, the optimal decision boundary discriminating R_M from R_N is not a plane and the linear discriminant function is inferior to a nonlinear one.

2.4 Nonlinear filtering with a neural network

In general, the probability functions of the signal one wants to identify and of the signals different to this are not Gaussian or have different covariance matrices. We propose to use nonlinear neural networks as alternative to a linear discriminant function. The intensities in the window $W(a,b)$ constitute the input to the network. Its output y is an estimate of the posterior probability that $W(a,b)$ contains the type of signal searched for, $\mathrm{p}(M|W(a,b))=\mathrm{P}(M)\ \mathrm{p}(W(a,b)|M)$ (see Eq. (3)). When $\mathrm{p}(M|W(a,b)) > 0,5$, the window $W(a,b)$ is most likely to contain a marker.

The neural network used in this paper is a multilayer perceptron with only one hidden layer with h nodes. It can be proved that such a nonlinear neural network can implement any arbitrary discriminant function when provided with a sufficient number of hidden nodes and trained with a large learning set.

Theorem 2.4.1. A multilayer perceptron with one layer of hidden nodes can implement an arbitrary discriminant function including a linear filter C.

Proof

It is proved in [1] that a multilayer perceptron with one layer of hidden nodes may implement any arbitrary discriminant function. ■

A neural network needs to be trained to implement the optimal discriminant function, the nonlinear filter that results in the best discrimination between markers and noise.

Theorem 2.4.2. The output y of a multilayer perceptron with one layer of hidden nodes will approach the posterior probability of the minimal error-rate classifier when the training set is large enough and the number of hidden nodes is sufficient.

Proof
It is proved in [2] that a multilayer perceptron with one hidden layer approaches the posterior probability of the minimal error-rate classifier when provided with a sufficient number of hidden nodes and trained with a large enough training set is used. ■

So a nonlinear neural network implements the optimal discriminant function when trained with a large set of images. In this paper, the training set is generated by a simulation model so the correct output value of the neural network is known (see appendix). Based on these theoretical results, we expect that a neural network will outperform a linear convolution filter for detection of markers.

3 Method

We suggest the following method to detection of markers. First, a neural network is used to classify a window $W(a, b)$ as containing a marker or not. When the output of the neural network $y \in (0, 1)$ exceeds 0,5 the neural network classifies the image as a marker, otherwise as containing noise. If $W(a, b)$ contains a marker, we will use a nonlinear regression neural network to estimate the central positions $(\hat{a},\hat{b})$ of the marker based on the intensities in $W(a, b)$.

4 Experiments

Two experiments with simulated markers were performed.

4.1 First experiment

In the first experiment, the performance of a neural network is compared with that of a linear filter. Both are trained to classify 500 images of the size 7×7 pixels, 250 images containing markers and 250 white noise (the appendix describes how the markers are generated). Neural networks with 49 input nodes, h hidden nodes, $h \in \{2, 3, 4\}$, and one output node were trained to detect whether the image contains a marker or not. The correct output of the neural network is $y = 1$ if the image contains a marker and $y = 0$ if not.

The networks were trained with back-propagation. Training continued until maximally 1 percent of the 500 training images were classified wrongly (marker classified as noise or the reverse). Table 1 contains the classification results obtained with a neural network and a linear filter.

4.2 Second experiment

In the second experiment, neural networks are trained to estimate the central position of a marker, $(\hat{a}, \hat{b})$, from the intensities in the window $W(a, b)$ containing a marker. Neural networks with 49 input nodes, h hidden nodes, $h \in \{6, 7\}$, and 2 output nodes were trained. In total 500 images containing markers were generated. The centre of each

Table 1 (left): Results from the first experiment. Detection of markers with a linear convolution filter and a neural network. Signal amplitude of marker=10-70 intensities, position=± 1 pixel, 500 test images, P(Marker)=0,5.

Correctness			
σ^2_{noise}	10	20	40
Linear filter	$0,7660$	$0,7253$	$0,7187$
NN, 2 hidden	$0,9273$	$0,8800$	$0,8527$
NN, 3 hidden	$0,9296$	$0,8771$	$0,8502$
NN, 4 hidden	$0,9282$	$0,8762$	$0,8482$

Standard dev. of estimated centre		
$\sigma^2_{noise} = 10$		
	σ_x	σ_y
NN, 6 hidden	$0,3372$	$0,3212$
NN, 7 hidden	$0,3255$	$0,3245$
$\sigma^2_{noise} = 20$		
	σ_x	σ_y
NN, 6 hidden	$0,3710$	$0,3409$
NN, 7 hidden	$0,3782$	$0,3379$

Table 2 (right): Results from the second experiment. Prediction of the centre of a marker. Signal amplitude of marker=10-70 intensities, position=± 2 pixels, 500 test images, the probability P(Marker) = 1.

marker is known with subpixel precision (see appendix). Each network was trained with back-propagation 10000 cycles to estimate the centre of the marker in each of the 500 images. The results of the second experiment are contained in table 2.

5 Discussion

The first experiment indicates that detection is performed better by the neural network than by the linear filter. This was to be expected as it was shown that linear filters are only optimal when the signals one wants to discriminate are normally distributed and the covariances of these two types of signals are identical. The neural network outperforms the linear filter for all three noise levels. The second experiment indicates that the central position of a marker can be estimated with a standard deviation of 0,32 to 0,38 pixel. Although encouraging, these results have not yet been compared with the performance of the standard linear method, the centre of mass.

Figure 1: Image containing 14 markers.

Appendix

The experiments are based images with artificial markers. A marker is a sphere with a known diameter d. In general, absorbtion is a function of the thickness of the material that is penetrated by an X-ray. As the diameter of the marker is known, its thickness t follows from

$$t = 2\,d\,\sqrt{l^2 - p^2} \tag{4}$$

with l the distance from the centre of the marker to the point through which the X-ray penetrates the sphere and p the projection of l onto the plane perpendicular to the penetrating X-ray. The absorbtion of an object is given by the equation

$$A(a,b) = A_0 \ \ \exp\left\{- \int_{c=0}^{t} \mu(a,b,c)\, dc\right\} \tag{5}$$

with (a,b) being the coordinates in a coordinate system that is perpendicular to the X-ray and μ the absorbtion constant of the material. The X-ray penetrates the object along the c-axis. The absorbtion of a marker has a circular symmetric form and its projection is a circle. For a pixel with the coordinates (e,f), when the centre $(a,b) = (0,0)$ the absorbtion can be computed from

$$A(a,b) = A_0 \ \ \exp\left\{- \ 2\ d\ \sqrt{(e+f)^2 - p^2}\right\} \tag{6}$$

with A_0 the intensity of the X-ray source and $d(=4)$ the diameter of the sphere in pixels.

The training sets used in both experiments consist of 500 images. With the probability 0,5 an image contains a marker which is generated using the absorbtion formula (6). The centre of the marker is chosen randomly from a uniform distribution, within ± 1 pixel (± 2 pixels in second experiment) from the centre (a,b) of the window. To centre the marker precisely on the randomly chosen centre within the image, subpixel translation in the Fourier space is used. The signal amplitude, the difference between the minimal and maximal intensity of the marker, is chosen randomly from a uniform distribution ranging from 10 to 70 intensities. Finally, Gaussian noise with a zero mean and a variance σ^2_{noise} is added to the marker. With a probability 0,5 the image should contain no marker and Gaussian noise is added with a zero mean and a variance σ^2_{noise}.

References

[1] Villiers, J.D. Bernard B.: Backprogation neural nets with one and two hidden layers, IEEE Transactions on Neural Networks, 4: 136-141, 1992.

[2] Richard, M.D. Lippmann, R.P.: Neural network classifiers estimate Bayesian a posteriori probabilities, Neural Computation, **3**: 461-483, 1991.

[3] Duda, R.O. Hart, P.E.: Pattern classification and scene analysis, John Wiley & Sons, New York, 1973.

[4] Niczyporuk, M.A. Craig Miller, D.: Automated tracking and digitization of multiple radio-opaque myocardial markers, Computers and biomedical research, 24: 129-142, 1991