Non-Parametric Image Subtraction using Grey Level Scattergrams

P. A. Bromiley, N.A. Thacker, P. Courtney

Imaging Science and Biomedical Engineering, Stopford Building, University of Manchester, Oxford Road, Manchester, M13 9PT.

Abstract

Image subtraction is used in many areas of machine vision to identify small changes between equivalent pairs of images. Often only a small subset of the differences will be of interest. In motion analysis only those differences caused by motion are important, and differences due to other sources only serve to complicate interpretation. Simple image subtraction detects all differences regardless of their source, and is therefore problematic to use. Superior techniques, analogous to standard statistical tests, can isolate localised differences due to motion from global differences due, for example, to illumination changes. Four such techniques are described. In particular, we introduce a new non-parametric statistical measure which allows a direct probabilistic interpretation of image differences. We expect this to be applicable to a wide range of image formation processes. Its application to medical images is discussed.

Key words: Image subtraction, grey-level scattergrams, multiple sclerosis.

1 Introduction

Image subtraction is a common tool for the analysis of change in pairs of images. The method is used in varying circumstances ranging from surveillance to interpretation of medical image data [7] [9] [6] [2]. Though this standard tool is quite simple, most researchers will already be familiar with the difficulties of interpreting the resulting difference image [4]. Taking a simple subtraction between two images and identifying regions of change using a threshold is directly equivalent to forming a null hypothesis test statistic, with the assumption of a single distribution for the expected level of change due only to

Email address: paul.bromiley@talk21.com (P. A. Bromiley).

noise that is the same across the entire image. In order for the technique to be used successfully great care has to be taken to ensure that the only changes taking place between the two images are due to the physical mechanisms that we are interested in. This may require that image data sets are realigned or pre-processed in order to remove gross changes before a subtraction can be performed. The result can always be used immediately to identify regions of maximal change, but ultimately we would also like to be able to put a quantitative statistical interpretation on the significance of the observed change. The problem that stands in the way of forming such an interpretation using conventional statistics is generally the lack of a known statistical model of the expected scene contents or perhaps even the imaging process. However, most images contain a sufficiently large quantity of data that in theory we might be able to extract sensible models of data behaviour from the data itself. This approach has been used widely in recent image registration techniques [11], particularly in medical applications [13]. The technique generally referred to as maximisation of mutual entropy is in fact a boot-strap approach to the construction of a maximum-likelihood statistic [8]. It therefore seems reasonable to attempt to adapt these measures, and equivalent approaches, to the problem of image subtraction in order to investigate the possibility of getting quantitative and statistically well-defined measures of difference for arbitrary image pairs.

2 Methods

The image subtraction techniques described here are based on the use of an image scattergram from a sample of image data as a basis for a statistical model. In order to construct the scattergram $S(g_1, g_2)$ from a pair of images, corresponding pixels were taken and their grey levels g_1, g_2 were used to define co-ordinates for entries in the scattergram. The two images are referred to below as the first and second images, and in each case the first image grey levels were plotted on the abscissa of the scattergram, and the second image grey levels on the ordinate. A vertical cut on the scattergram $F(g_1 = const, g_2)$ isolates a set of pixels in the first image with the same grey level. The distribution of this data then gives the relative frequency of grey levels for pixels in the second image.

Iterated tangential smoothing was applied to the scattergram to ensure that the surface in grey-level space was smooth and continuous. This technique applies local averaging of three grey level values lying on the tangent to the local direction of maximum gradient in the scattergram. By definition there should be zero derivative in the tangential direction for a 2D data set. This technique therefore smooths the surface in grey-level space in a way that preserves the original data distribution. It can be shown using the central limit theorem

that an n-iterated tangential smoothing approximates an oriented Gaussian filter of width $\sigma = \sqrt{n/2}$. The main advantage of tangential smoothing over alternatives such as Gaussian smoothing is that the former does not increase the overall size of features in the grey level space, an effect that was found to be detrimental to the performance of some of the subtraction algorithms described below.

For similar images, the ridge in the scattergram provides a model describing the global variations between the two images. For instance, if the level of illumination in the second image was increased, the ridge would move upwards in the scattergram. The ridge can therefore be used to differentiate between global differences due to illumination changes and localised differences caused by motion. Four image subtraction methods based on image scattergrams were investigated, and each of these is described in detail below. Fig. 1 shows a schematic diagram of the new image subtraction methods.

2.1 Spline Ridges

The simplest method chosen for extracting the position of the ridge in the scattergram was to fit it over a series of ranges with polynomials. Cubic splines were used to guarantee that the curve was smooth and continuous at the boundaries between the ranges. The splines were fitted to a series of knot points on the ridge. Initial approximations for these knot points were chosen by specifying the number of points n desired, dividing the graph into $n-1$ ranges in the horizontal direction, and searching along the boundaries of the ranges in the vertical direction to find the maximum values.

Once the initial approximations to the knot points had been found, simplex minimisation was used to locate the ridge [5]. The cost function for the spline fit was defined as the negative sum of the grey levels of the pixels i lying under the spline curve $C(g_1, \mathbf{g_2^{max}})$ in the scattergram.

$$Cost \;=\; -\;\sum_i F(g_i, C(g_i, \mathbf{g_2^{max}}))$$

Once the spline fit had been optimised in this way, it was used to produce a new version of the second image. The grey levels of pixels from the first image were used to define a vertical cut on the scattergram, and the corresponding vertical ordinates for the ridge in the scattergram were found using the spline curves. The vertical ordinates were then used as grey levels in the new second image, which could then be subtracted from the original second image to give a difference image.

$$D_1(x,y) \;=\; I_2(x,y) \;-\; C(I_1(x,y), \mathbf{g_2^{max}})$$

The new second image is an estimate of the first image, scaled to remove the global variations between the original images encoded in the ridge in the scattergram. The subtraction of the new second image from the original second image therefore effectively removed any global mapping between grey level values in the two images.

2.2 Log-Likelihood

An alternative to the splines-based method for calculating the difference image that did not rely on fitting curves or finding the maxima in the scattergram was also implemented. Normalising the profile along the vertical cut $g_1 = k$ gave a probability distribution, with each graph pixel corresponding to the probability of obtaining a particular grey level value at a pixel in the second image given the grey level value of the same pixel in the first image. The difference image was then produced by taking corresponding pairs of pixel grey levels and producing a difference image with values given by -2 times the natural log of the scattergram value.

$$D_2(x,y) = -2\ ln[\frac{F(I_1(x,y),I_2(x,y))}{\sum_c F(I_1(x,y),c)}]$$

Assuming that the probability distribution takes the form of a Gaussian,

$$P(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}},$$

where P is the probability of obtaining a measurement x from a Gaussian distribution of mean μ and width σ, then the $-2lnP$ is given by

$$-2lnP = 2ln\sigma\sqrt{2\pi} + \frac{(x-\mu)^2}{\sigma^2}.$$

Comparing this to the standard form for a z-score, $(z = (\frac{x-\mu}{\sigma}))$, it can be seen that the right-hand side of the equation for $-2lnP$, ignoring the first term, is equivalent to the square of the z-score. The advantage here however is that the scattergram will model any arbitrary distribution, not just a Gaussian, and this also includes distributions with arbitrary width.

This difference image, when summed, gives an overall difference statistic which is directly equivalent to the mutual entropy measures used for image co-registration. We may therefore expect that the returned values can be of use in terms of arriving at sensible statistical decisions regarding the level of difference between the two images. Unfortunately, this is only the case when the

spread of measured values is the same for all mean values. We must therefore seek another measure if we wish to make statistical decisions regarding individual pixels.

2.3 Local Maxima

In some cases we may be given pairs of images for which there are ambiguous grey level regions. This may manifest itself as bi-modal distributions within the scattergram. One way to deal with this is to take the difference not from the global maximum along a cut, but from the nearest local maxima. In order to calculate the z-score with reference to the nearest local maximum the peaks in the scattergram were located by a simple search. The grey levels for a pair of pixels from the original images defined co-ordinates in the scattergram that acted as the starting point for the search. The search then proceeded both upwards and downwards in $F(g_1, g_2)$. The peaks were roughly located, and then the position was refined by interpolation using a quadratic fit to the three points around the peak. This gave the vertical ordinate g_{cpeak} of the nearest peak in the scattergram, which was subtracted from the vertical ordinate of the starting position to give an effective z-score. The difference image can be defined as

$$D_3(x, y) \;=\; I_2(x, y) \;-\; g_{cpeak}$$

The z-score was used as the grey level in the difference image, and the procedure was repeated for every pair of pixels from the original images.

2.4 Probability Integration

The final method we devised for calculating the difference image used the probability distributions in the scattergram directly, without calculating an effective z-score. The idea was to try to construct a probability value that reflected how likely it was that each grey level had been drawn from the same generation process as the rest of the data. A vertical cut in the normalised scattergram gave a probability distribution describing the grey levels of a set of pixels in the second image that all had the same grey level in the first image. As in the previous methods, the grey levels of a pair of pixels from the original images were used to define a set of co-ordinates in the scattergram. An integration was then performed along the vertical cut passing through this point, summing all of the values **smaller** than that of this pixel. This total

was used as the grey level value for the relevant pixel in the difference image [1].

$$D_4(x,y) \;=\; \sum_c \delta(\; F(x,c) > F(x,y) \;)F(x,c)$$

This technique produces a difference image in which the grey level of each pixel is the probability of the pairing of grey levels for the corresponding pixels in the original images. The distribution of grey levels in the difference image is by definition flat for data drawn from the mean distribution. Such probability distributions are **honest** [3], i.e. a 1% probability implies that data will be generated worse than this only 1/100 th of the time. The measure has the same interpretation as the conventional "chi-squared probability" except that, as we do not need to specify a particular distribution, it is essentially non-parametric. Low probabilities indicate that the pairings of pixels are expected to be uncommon. This is exactly the type of measure we need in order to identify outlying combinations of pixel values in an automatic manner, solving the problems inherent in the likelihood-based approach.

The distribution of grey levels in the difference images produced using this method can act as a self-test: ignoring the low-probability pixels generated by localised differences, it should be flat. Any significant departure from a flat distribution therefore indicates inappropriate behaviour of the two data sets and therefore unsuitability of the statistic for that comparison.

3 Results

3.1 Testing with Synthetic Data

The four routines described above were tested using synthetic data to ensure that they were working as predicted. They were designed to be able to ignore effects such as changes in the shading or illumination of features, and focus only on movement of features between images. Two test images were prepared using the image creation tool in TINA [10]. Each was an 8bpp greyscale image, 256 by 256 pixels in size, of a 128 by 128 pixel rectangle in the centre of the frame. The positions of the rectangles in the two frames were offset by one pixel to simulate the effects of motion. The rectangles were shaded so that their grey levels varied smoothly between 30 on one vertical edge and 200 on the other, with the shading being uniform in the vertical direction. The direction of shading was reversed between the two images. Finally, Gaussian noise with a standard deviation of 10 grey levels was added to both images. The

[1] δ represents the Kronecker delta function.

resultant test images are shown in Fig. 2, together with the scattergram and the results of a simple pixel-by-pixel subtraction. It was anticipated that the image subtraction techniques described above should be capable of detecting the one-pixel shift in the position of the rectangles whilst ignoring the difference in shading, and return difference images containing two one-pixel wide vertical lines at the positions of the vertical edges of the rectangles, against a background uniform random noise. The difference images produced by the four methods are shown in Fig. 3.

The splines-based method shows the greatest departure from the expected results, and this was due to a failure in the spline fitting. The fit was good for the peak close to $(0, 0)$, which corresponded to the background, and close to the central part of the scattergram. However, it failed in the regions corresponding to the highest and lowest grey level values in the rectangle in the first image, due to the discontinuities in the scattergram at these points. Therefore, the difference image departs from the expected output around the vertical edges of the rectangle.

The difference images generated by the remaining three methods all closely match the expected output. The vertical features at the vertical edges of the rectangles show the capability of these methods for detecting movement of image features whilst ignoring other effects. The difference image returned by the probability integration technique has a flat probability distribution and so appears noisier than the other difference images. This is in fact the main advantage of the technique, since it allows subsequent quantitative processing of the difference image.

3.2 Testing with Image Sequences

In order to convincingly demonstrate the ability of these methods to detect movement in images, they were applied to a pair of images from a sequence showing a moving train. The original images, the scattergram, and the results of a simple subtraction are shown in Fig. 4. The four difference images generated using the new methods are shown in Fig. 5.

In order to show that the new image subtraction techniques were capable of detecting any differences due to motion whilst ignoring those due to illumination changes, an artificial illumination change was applied to the second frame in the form of a multiplicative constant. The effects of this illumination change can be seen clearly in the scattergram. Since the differences due to motion between the two images are small, almost all the points in the scattergram lie on a straight line. In the absence of any illumination difference, they would lie along the line $x = y$. In this case, the gradient of the line is considerably

steeper.

The results of the simple subtraction show all of the differences between the two images, and so the small differences due to motion are hidden by the larger differences due to illumination. However, the new techniques ignored the illumination differences, and so the resulting difference images show only the differences due to motion at the boundaries of the moving train. The tracks and the ruler next to them were also detected to varying extents, since the camera was focussed on the front of the train, and so the point of focus shifted as the train moved towards the camera. The consequent changes in the degree of blurring at any given point on these features were greatest close to the point at which the camera was focussed. The region of the ruler closest to the camera was saturated in both of the original images, suppressing the noise, and this was highlighted by the probability integration technique.

3.3 Testing with Medical Data

The image subtraction techniques described here can also be applied to medical image data. For example, MS lesions in the brain can be difficult to detect in an MRI scan, but can be highlighted using an injection of gadolinium (GdDTPA), which concentrates at the lesion sites. Scans taken before and after the injection can be subtracted to help identify lesions, but the gadolinium also alters the global characteristics of the scan, so a simple pixel-by-pixel subtraction will not remove all of the underlying structure of the brain from the image. The new image subtraction methods should be able to take the global changes into account, and thus produce an image which shows only the lesions.

Obtaining a gold standard for this work is difficult without extensive hystological investigation. In order to simulate the imaging process, two T2 scans with slightly different echo train times (TE) of the same slice of a brain were used. This simulates the effects of repeat scanning on different scanners after a significant time interval and the small quantitative changes which occur in signal due to the presence of a contrast agent. The background was removed from the image so that our statistical model (scattergram) was estimated using only the tissues in which we were interested. It was expected that, under many circumstances, corrections for intensity non-uniformities in the data would need to be applied [12], but this was not found to be neccessary with these images. A grey level offset too small to be detected visually was then added to a small circular region of one of the brain images. This simulates the typical changes in recovered signal intensity which would occur due to lesions but in a testable manner. The magnitude of the offset was based on the noise in the original images, calculated from the width of vertical and horizontal

gradient histograms of the original images around zero. The subtraction routines were applied to this data in an attempt to detect the change. Fig. 6 shows the brain images with an offset of 2σ added to a small region of one image, together with the scattergram and the results of a simple subtraction. The altered region cannot be detected visually in the original images, and is barely visible in the pixel-by-pixel difference image. Fig. 7 shows the difference images generated using the methods described above, and the altered region shows up clearly in the output from the log-probability and integration-based methods. The altered region ceased to be detectable when the magnitude of the offset was reduced below around 1σ. Fig. 8 shows a histogram generated from the integration-based difference image demonstrating that, as expected, this method produced an honest probability distribution, confirming that this statistical measure is appropriate for this MR data.

Finally, in order to demonstrate the use of the new subtraction techniques on genuine MRI data, they were applied to a pair of scans of the brain of an MS patient, taken before and after the injection of the GdDTPA contrast agent used to highlight lesions in the brain. Fig. 9 shows the two scans, together with the scattergram and the results of a simple subtraction. The larger lesions can easily be seen in both the second scan, as the lighter regions, and the reults of the simple subtraction, as the darker regions. The results of the new image subtraction methods are shown in Fig 10. As in the previous examples, the nearest peak based method allowed the implicit model to match all of the data in the images, and so did not shows the lesions in the output images. The splines-based method also clearly shows the lesions, but again fitting failures result in this method retaining a large amount of the underlying structure of the brain in the output image. The log-probability method shows a considerable improvement in this respect and again clearly shows the lesions as lighter regions. Finally, the intergration based method removed all of the underlying brain structure from the image, leaving only the lesions as darker regions against a background of random noise with a flat probability distribution. The lesion sites detected by the new image subtraction techniques matched those identified by the radiologist.

In these images the lesions were easily visible, and so visual inspection of the simple subtraction results by the radiologist was sufficient to identify the lesion sites. In this case the main advantages of the new subtraction techniques, particularly the integration-based technique, was that it produced results in terms of a well-defined statistical quantity. This would allow further processing of the images to be conducted in a quantitative manner. For instance, Fig. 11 shows the results of thresholding the integration-based difference image at the 1% level. The integration-based method produces an output image with a flat probability distribution for pixels drawn from the mean distribution, and so thresholding the image at the 1% level will extract the low-probability pixels not drawn from the mean distribution i.e. those corresponding to le-

sions, together with 1% of the pixels from the normal brain tissue. Since the proprotion of normal tissue pixels extracted is known, volumetric analysis of the lesions can now be performed using thresholding. The issues of enhancing the contrast of MS lesions in MRI scans and of volumetric analysis of the lesions are important areas, in relation to both tracking the progression of the subclinical disease, and to theraputic trials [14].

4 Discussion and Conclusions

Simple pixel-by-pixel image subtraction, when considered as a statistical test, can be shown to be making many assumptions regarding the information contained in a pair of images. These assumptions are rarely valid and as a consequence simple image subtraction cannot be used reliably [4]. However, there is often considerable information available within a pair of images, which can be used to avoid having to make all of these assumptions. The new image subtraction techniques described here used a scattergram of the grey levels in a pair of images to extract this information. The scattergram is effectively used as a model of the global variations between the images, allowing the new techniques to focus only on the localised variations.

The splines-based method was the simplest of the four new techniques, but also the most inaccurate, due to failures in the spline to model the ridge in the scattergram. In practice the technique will always be susceptible to this effect, as there is no fitting method that can guarantee to fit to any arbitrary data set. The main result for this method was therefore to demonstrate that a more sophisticated technique was needed (though the method may have some merit on small datasets).

An attempt to model bi-modal distributions seemed to give results which were not particularly informative. Selection of the nearest peak simply allowed all data to be reasonably consistent with the model.

The capabilities of the log-probability and integration-based techniques were clearly demonstrated using the synthetic data. The images of the moving train also clearly showed both the ability of the new methods to extract the boundaries of a moving object whilst ignoring differences due to global illumination changes, and their advantages over simple subtraction. When both sets of results are considered together, it is clear that the new techniques will be superior for motion detection in more realistic environments. These techniques were also shown to be superior in detecting abnormalities in medical images, demonstrating the wide range of potential applications.

One of the most important features of the log-probability and integration-

based methods is that the grey levels in the difference images correspond to well-defined quantities: the square of the z-score and a probability respectively. This is in sharp contrast to pixel-by-pixel image subtraction, where the grey levels in the difference image are arbitrary measures of difference in units of grey levels, and have no objective meaning.

The log-probability and integration-based techniques can be considered as the definitions of new non-parametric statistical tests, with theoretically predictable properties. As such, these methods are firmly grounded in the existing body of statistical decision theory and can be used in combination with methods designed for more restrictive parametric techniques for the testing of hypotheses. This fact makes extensive quantiative analysis redundant, though the results presented here already demonstrate the applicability of these measures to motion analysis and subtraction of MR datasets.

The integration-based method produces an output image where the probability distribution is flat for pixels drawn from the mean distribution, which provides a mechanism for self-test. In addition, such statistical methods permit data interpretation using only the single model of interest. In contrast, Bayesian techniques for computing the probability $P(C_i|D)$ that a given model C_i explains the data D require prior knowledge of models representing all the distributions present in the data,

$$P(C_i|D) = \frac{P(D|C_i)}{\sum_i P(D|C_i)}$$

This difference can be important in some data interpretation tasks.

A standard technique exists for combining and re-flattening independent flat probability distributions [1]. If n probabilities P drawn from a flat distribution are multiplied to produce a combined probability,

$$P = \prod_i^n P_n$$

then the combined probability can be re-flattened using

$$P' = P \sum_{j=0}^{n-1} \frac{(-lnP)^j}{j!}$$

This process is potentially nestable, allowing region-based data fusion to produce a principled statistical test of whether the data are drawn from the mean distribution. These measures could also be used in the analysis of co-occurence

of spatially distributed values and may thus also have a role in the analysis of texture. This is an area that we now intend to investigate.

References

[1] ALEPH Collaboration, A Precise Measurement of $\Gamma_{Z\to b\bar{b}}/\Gamma_{Z\to hadrons}$, Phys. Lett. B313(1993) 535.

[2] Baumberg A, and Hogg D, Learning Flexible Models from Image Sequences, Proc. ECCV 1994, J-O. Ekhlund (Ed), Stockholm, (1994) 299-308.

[3] Dawid A P, Probability Forecasting, in: Encyclopedia of Statistical Science 7, (Wiley, 1986), 210-218.

[4] Hajnal J V, Young I R, and Bydder G M, Contrast Mechanisms, in: Bradley Jr W J, and Bydder G M, eds. Functional MRI of the Brain, Advanced MR Imaging Techniques, (Martin Dunitz Ltd, London, 1997) 195-207.

[5] Press W H, Teukolsky S A, Vetterling W T, and Flannery B P, Numerical Recipies in C 2nd Ed., (Cambridge University Press, 1992).

[6] Koller D, Weber J, and Malik J, Robust Multiple Car Tracking with Occlusion Reasoning, Proc. ECCV 1994, J-O. Ekhlund (Ed), Stockholm, (1994) 189-196.

[7] Murray D, and Basu A, Motion Tracking with an Active Camera, IEEE Trans. Pattern Analysis and Machine Intell., 16(5), (1994) 449-459.

[8] Roche A, Malandain G, Ayache N, and Prima S, Towards a Better Comprehension of Similarity Measures Used in Medical Image Registration. Proc. MICCAI, 99. Cambridge, (1999) 555-565.

[9] Rowe S, and Blake A, Statistical Mosaics for Tracking, Image and Vision Computing, 14(8), (1996) 549-564.

[10] Thacker N A, Lacey A, Vokurka E, Zhu X P, Li K L, and Jackson A, TINA an Image Analysis and Computer Vision Application for Medical Imaging Research. Proc. ECR, s566, Vienna, 1999.

[11] Viola P, Alignment by Maximisation of Mutual Information. PhD. Thesis, MIT, Artificial Intelligence laboratory, 1995.

[12] Vokurka E, Thacker N A, and Jackson A, A Fast Model Independant Method for Automatic Correction of Intensity Non-Uniformity in MRI Data, JMRI 10(4) (1999) 550-562.

[13] West J, et al. Comparison and Evaluation of Retrospective Intermodality Image Registration Techniques. J.Comp.Assist.Tom 21(4) (1997) 554-566.

[14] Wolansky L J, et al. Triple-Dose Versus Single Dose Gadoteridol in Multiple Sclerosis Patients, Journal of Neuroimaging 4(3) (1994) 141-145.

5 Acknowledgements

The authors would like to acknowledge the support of the MEDLINK programme, grant no. P169, in funding part of this work. All software is freely avalable from the TINA website www.niac.man.ac.uk/Tina.

6 Figures

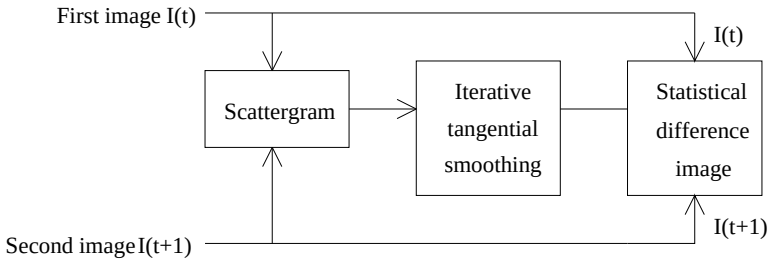


Fig. 1. The operation of the new image subtraction routines.

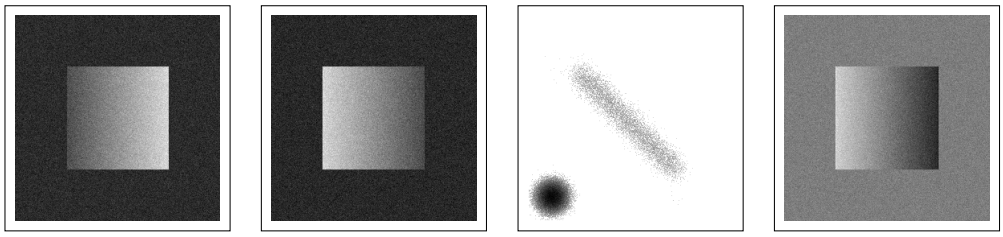


Fig. 2. The original images (a,b), scattergram (c), and simple subtraction difference image (d) for the synthetic data.

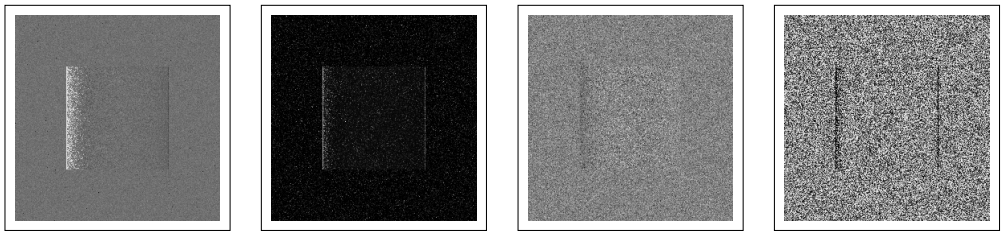
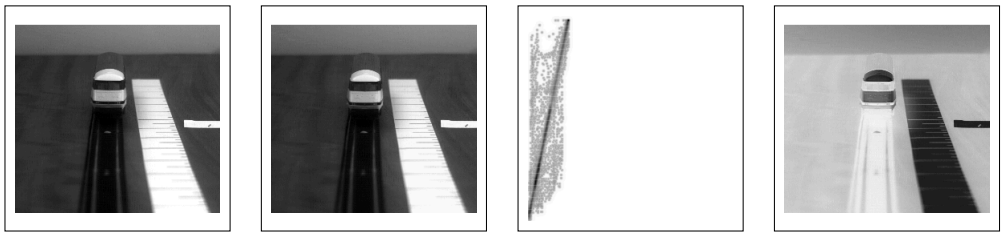
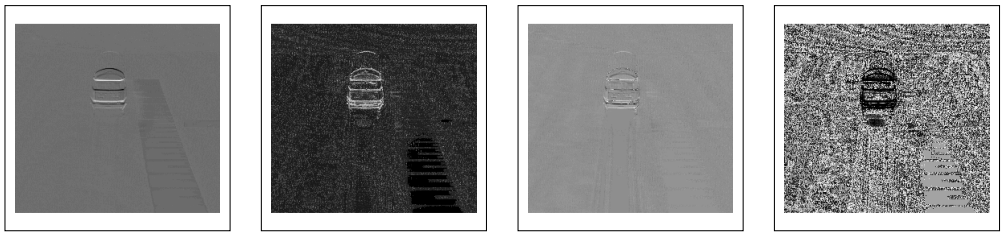


Fig. 3. The splines-based (a), log-likelihood (b), nearest peak (c), and integration-based (d) difference images for the synthetic data.



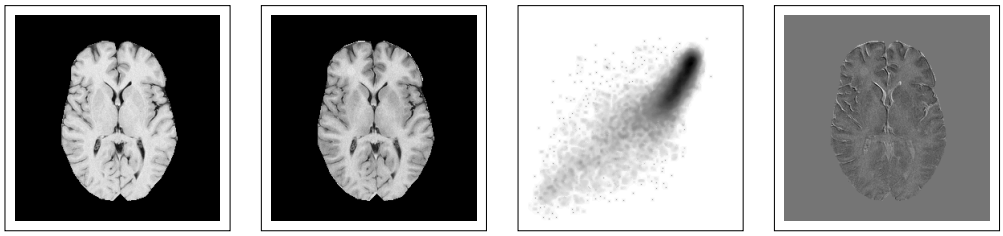
(a) (b) (c) (d)

Fig. 4. The original images (a,b), scattergram (c), and simple subtraction difference image (d) for the train sequence.



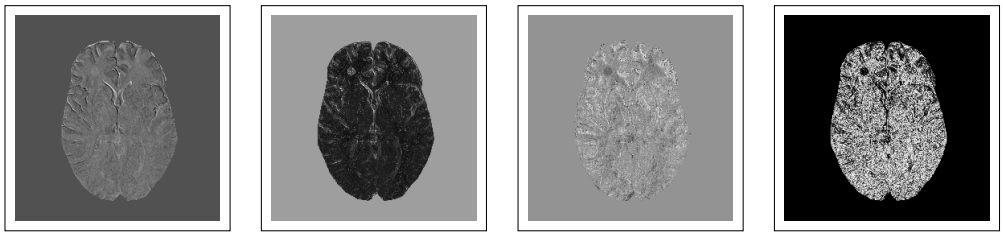
(a) (b) (c) (d)

Fig. 5. The splines-based (a), log-likelihood (b), nearest peak (c), and integration-based (d) difference images for the train sequence. The apparent extra noise in the integration-based difference image is in fact a consequence of its flat probability distribution.



(a) (b) (c) (d)

Fig. 6. The original MRI brain images (a,b), scattergram (c), and simple subtraction difference image (d), with a 2σ offset added to a small region of image (b).



(a) (b) (c) (d)

Fig. 7. The splines-based (a), log-likelihood (b), nearest peak (c), and integration-based (d) difference images for the MRI brain scan images. The log-likelihood and integration-based difference images clearly show the altered region in the upper-left area of the image.

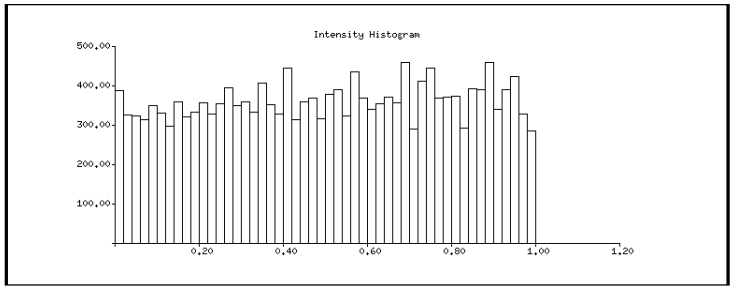
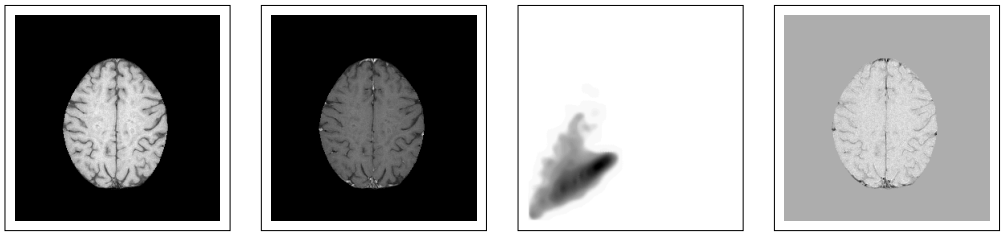
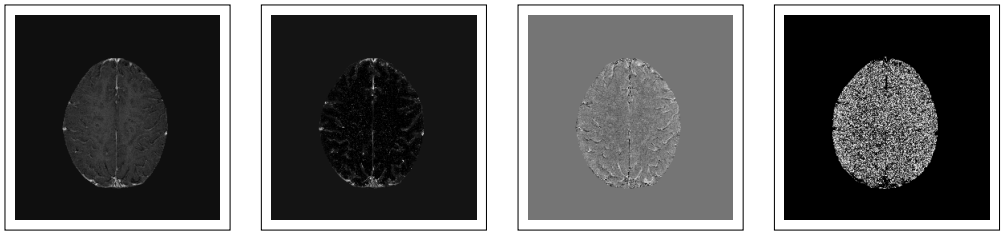


Fig. 8. Histogram of the integration-based difference image generated from the MRI brain images, showing the flat probability distribution.



(a) (b) (c) (d)

Fig. 9. MRI brain scans from an MS patient before (a) and after (b) a GdDTPA injection, with the scattergram (c) and simple subtraction difference image (d).



(a) (b) (c) (d)

Fig. 10. The splines-based (a), log-likelihood (b), nearest peak (c), and integration-based (d) difference images for the MRI brain scans from an MS patient.

Fig. 11. Results of thresholding the integration-based subtraction at the 1% level.