Magnetic Resonance Image Processing Using Levy Distributed Anisotropic Diffusion

Jonathan Blackledge
Technological University Dublin, jonathan.blackledge59@gmail.com

Matthew Blackledge
Institute of Cancer Research, London

Follow this and additional works at: https://arrow.dit.ie/engscheleart
Part of the Electrical and Computer Engineering Commons

Recommended Citation
Magnetic Resonance Image Processing using Lévy Distributed Anisotropic Diffusion

Jonathan Blackledge† and Matthew Blackledge∗

School of Electrical Engineering Systems
Dublin Institute of Technology
Dublin, Ireland

EPSRC Cancer Imaging Centre
Institute of Cancer Research
London, England

E-mail: †jonathan.blackledge@dit.ie
∗mattyblackledge@gmail.com

Abstract — We consider the physical nature of the self-diffusion of water molecules in tissue and explore how (Nuclear) Magnetic Resonance (MR) imaging may be used as a means of measuring the rate of diffusion in vivo. A discussion is presented on how these techniques may be implemented as a non-invasive means of assessing the response of tumours to novel therapeutics including some of the basic advantages and disadvantages when compared to other methods. The physical basis and mathematical models for diffusion are considered together with models for the distribution of the diffusion coefficient including a Lévy distributed model. Using a Lévy distributed diffusion model, we develop a novel algorithm for the purpose of improving the signal-to-noise ratio of MR images.

Keywords — Magnetic Resonance Imaging, Diffusion Imaging, Fractional Diffusion, Noise Reduction.

I Introduction

Investigations into the use of NMR measurements to study the self-diffusion of water began with work by Hahn in 1950 who discovered that there was an inherent loss in signal in spin-echo sequences due to the motion of water during the application of magnetic gradients[1]. These ideas were later developed by Stejskal and Tanner [7] who showed that using two gradients of opposite polarity either side of the π rf excitation pulse in spin-echo sequences resulted in signals whose amplitudes were highly dependent on the diffusivity of the medium being probed. Figure 1 shows a simplified Stejskal-Tanner MRI pulse sequence (also known as the pulsed field gradient or PFG sequence) in which the diffusion sensitizing gradients have been labelled with a duration of δ ms, a gradient magnitude of g Tm⁻¹ and the second pulse has been applied at a time of Δ ms after the first (Δ/2 ms after the π rf pulse).

Let us consider the effect of this pulse sequence on a single spin at position r. Firstly the application of a π/2 rf pulse flips the spins’ magnetization
vector, \( \mathbf{m} \), into the \( x'-y' \) plane of the rotating frame of reference (\( z' \) is taken to lie along the same direction as the main \( B_0 \) field). Following this the first diffusion sensitizing gradient (\( g_1 \)) rotates \( \mathbf{m} \) by an angle of \( \alpha_1 = \delta y \mathbf{r} \cdot g_1 \) about \( z' \) and then a \( \tau \) rf pulse flips \( \mathbf{m} \) by 180° about either \( x' \) or \( y' \) depending on the choice of the pulse direction. Finally, the second diffusion gradient (\( g_2 \)) rotates \( \mathbf{m} \) through an angle of \( \alpha_2 = \delta y \mathbf{r} \cdot g_2 \) about \( z' \). Provided that \( g_1 = g_2 \) and that the spin does not move during the pulse sequence then the net phase offset acquired (neglecting \( T_2 \) effects) is \( \alpha_1 - \alpha_2 = 0 \). However if we were to consider that the spin has time dependent position, \( \mathbf{r}(t) \), then the effect of the diffusion pulses would be to rotate the \( \mathbf{m} \) through a total angle of

\[
\theta = \int_0^\delta \gamma \mathbf{r}(t) \cdot g dt - \int_\Delta^{\Delta + \delta} \gamma \mathbf{r}(t) \cdot g dt
\]

around \( z' \). By considering that the spins are randomly moving due to diffusion, then the spins which form an isochromat at a position \( \mathbf{r}_0 \) at the time of signal acquisition will have acquired a distribution of phase offsets during the sequence depending on the diffusivity of the material. In a similar fashion to \( T_2 \) dephasing the bulk magnetization (and hence the NMR signal) will have a smaller magnitude than if no diffusion were present due to the ‘fanning out’ of spins (see fig. 1) .

A mathematical framework for these ideas can be built using complex notation where a phase shift \( \theta \) may be represented by multiplying the original \( \mathbf{m} \) by \( e^{i \theta} \). Therefore the net contribution of a cohort of spins at position \( \mathbf{r}_0 \) on the signal will be [5]

\[
S(\tau) = S(\tau)_{g=0} \int_{-\infty}^{\infty} P(\theta, \Delta) e^{i \theta} d\theta
\]

where \( S(\tau)_{g=0} \) is the signal which would have been recorded had the diffusion gradients been omitted and \( P(\theta, \Delta) \) is the normalized probability of finding a spin with phase shift \( \theta \) at \( \mathbf{r}_0 \) acquired during the time period \( \Delta \) (see figure 1). As discussed before, it may be shown that in the case of diffusion the integral results attenuation of the signal amplitude. It is therefore sensible to write this equation in the form

\[
\frac{|S(\tau)|}{|S(\tau)_{g=0}|} = E(\delta, \Delta, g, D)
\]

where \( E \) is the attenuation of the signal due to the parameters of the PFG sequence and the diffusion coefficient of the medium. The question then remains of finding a numerical solution for \( E \) so that the signal attenuation is quantifiable in terms of the diffusion coefficient and the PFG pulse sequence parameters. One method of solution, known as the ‘Gaussian Phase Distribution’ (GPD) approximation, involves the assumption that the form of \( P(\theta, \Delta) \) is equivalent to the probability distribution of the distance travelled by a particle in time \( t \) undergoing unrestricted isotropic diffusion so that

\[
P(\theta, \Delta) = (2\pi\langle \theta^2 \rangle)^{-\frac{1}{2}} \exp\left(-\frac{\theta^2}{2\langle \theta^2 \rangle}\right)
\]

where \( \langle \theta^2 \rangle \) is the expected square phase shift, found by evaluating equation (1) with \( \langle \mathbf{r}^2 \rangle = \langle (\mathbf{r}_1 - \mathbf{r}_0)^2 \rangle = nDt \) giving (as derived in detail by Price [5])

\[
\langle \theta^2 \rangle = \gamma^2 |g|^2 n D \delta^2 \left(\Delta - \frac{\delta}{3}\right)
\]

where \( n = 2, 4, 6 \) for 1, 2 and 3 dimensions respectively (\( n = 2N \)). In the case of DWI \( g \) is only applied in one direction at any given time so it is sufficient to consider the effects of diffusion in only one dimension where \( N = 1 \). By substituting \( P(\theta, \Delta) \) into equation (2) and using the standard integral [2]

\[
\int_{-\infty}^{\infty} e^{-p^2 x^2} \pm q x dx = \frac{\sqrt{\pi}}{p} e^{-\frac{q^2}{4p^2}}
\]

it may be shown that

\[
E(\delta, \Delta, g, D) = \exp\left(-\frac{\langle \theta^2 \rangle}{2}\right)
\]

\[
= \exp\left[-\gamma^2 |g|^2 D \delta^2 \left(\Delta - \frac{\delta}{3}\right)\right]
\]

This formula is usually simplified by combining all the scanning parameters into a single variable, \( b \), giving

\[
E(b, D) = e^{-bD}, \quad b = \gamma^2 |g|^2 D \delta^2 \left(\Delta - \frac{\delta}{3}\right)
\]

As the scanning parameters (and hence \( b \)) may be chosen at will (within certain limits depending on the instrument) it is possible to obtain different levels of contrast in MR images that are dependent on the diffusivity of the medium. Furthermore, if a number of signals are acquired with different \( b \)-values (at least 2) then the gradient of the log plot of signal versus \( b \) gives an estimate for the value of the diffusion coefficient.

It is important to note that in the above arguments the assumptions are made that (a) diffusion is isotropic, \( b \) diffusion is unrestricted and (c) there is no net flow in the probed sample of water. Clearly in all situations assumption (b) is going
to be violated but provided that the encasement of the water is relatively large and $\Delta$ is kept relatively short then unrestricted diffusion is a good approximation. Assumption (a) also holds under these conditions as the effects of anisotropy mainly come into play when considering spins that heavily restricted in one direction and not in another. However, in the case of imaging in vivo as will be explained later the assumptions do not hold well due to the restriction of water molecules by a complex entanglement of cell walls. It should be noted that other solutions of $P(\theta, \Delta)$ exist (see Price [5]) depending on which assumptions seem most valid for the given system but they are beyond the context of this paper.

II Biophysics of Diffusion

Whilst the above discussion provides insight into how diffusion is measured through magnetic resonance techniques, a number of simplifying assumptions have been made which are not valid in biological systems. Firstly, in the derivation of the GPD approximation it is assumed that the diffusion of water molecules is free. However, in tissue diffusivity measurements the motion of water molecules is hindered and restricted by the various cellular constituents and cell walls [3] and so it is common to use the term apparent diffusion coefficient (ADC) [6]. Furthermore, most tissues contain of a complex array of capillaries in which the effects of flow and perfusion must also be considered. Due to their small length scales and the fact that they are generally quite tortuous (especially in the case of tumour tissues) the movement of water within the capillary network may be viewed as a random walk and hence is similar to that of diffusion in the sense discussed above. However, as the motion is generally quicker due to flow the pseudo-diffusion coefficient, $D_{\text{p}}$, for water within the capillaries is likely to be higher than that for water in which no flow is present [4]. From these arguments it is possible to arrive at the Intravoxel Incoherent Motion (IVIM) biexponential relation where the total signal attenuation is proposed to consist of the sum of the attenuation due water within capillaries and the attenuation due to water in the rest of the tissue, each multiplied by the fraction of the voxel volume they occupy:

$$E(b) = \frac{S(b)}{S(b = 0)} = \alpha e^{-bD_{\text{p}}} + (1 - \alpha) e^{-b\text{ADC}}$$

Assuming that $D_{\text{p}} > \text{ADC}$ then for small values of $b$, $E$ will be dominated by the attenuation due to pseudo-diffusion and similarly for high $b$ values the reverse is true. For this reason measurements of the diffusion coefficient are often made over low and high $b$-values and are named ADC\text{fast} and ADC\text{slow} which are estimates for $D_{\text{p}}$ and $\text{ADC}$ respectively. Whilst these ideas help in understanding the physical properties of the tissue, the use of only two diffusion coefficients is still rather limited. It is perhaps more appropriate to consider that the tissue has some distribution of possible $D$ values all of which contribute to the observed signal attenuation. The ideas behind the biexponential model can be generalized by:

$$E(b) = \sum_{i=0}^{\infty} C_i e^{-bD_i}$$

where $C_i$ is the voxel fraction of tissues with diffusion coefficient $D_i$. If $C_i$ is then considered to be a continuous function $P(D)$ which is normalised to unity, then the signal attenuation is described as

$$E(b) = \int_{0}^{\infty} P(D) e^{-bD} dD = \mathcal{L}\{P(D)\}$$

where $\mathcal{L}$ represents the Laplace transform. An ideal experiment would be to calculate the inverse Laplace transform of the signal decay profile obtained at many $b$-values to give diffusion coefficient probability density function. However, in practice this is extremely difficult as a small change in $E(b)$ can be the cause of a very large change in $P(D)$. It would therefore be necessary to calculate the attenuation curve over a large range of many closely spaced $b$-values and the signal to noise ratio would need to be exceptional. An alternative approach is to hypothesize a distribution $P(D)$, which could be done theoretically or perhaps through Monte Carlo simulations, and consider the Laplace transform as fitting function for $E(b)$. As an example one may assume that $P(D)$ is normally distributed such that the mean diffusion coefficient, $\bar{D}$, has the highest voxel fraction and the width of the curve is described by the standard deviation of the coefficients, $\sigma$:

$$P(D) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left\{ -\frac{(D - \bar{D})^2}{2\sigma^2} \right\}$$

![Fig. 2: A simplified diagram showing the typical structure of cells and capillaries in tissues.](image)
Analyzing the Laplace transform yields

\[ E(b) = \frac{1}{2} \exp \left\{ -b\bar{D} - \frac{1}{2} b^2 \sigma^2 \right\} \times \text{erfc} \left\{ \frac{1}{\sqrt{2}} \left( b\sigma - \frac{\bar{D}}{\sigma} \right) \right\} \]

However, one major shortcoming of this distribution is that it allows for negative \( D \) values which are clearly unachievable in reality. Furthermore, it is likely that distribution would have some skewness and kurtosis which cannot be described using standard normal distributions. Another approach is to use a function which best fits \( E(b) \) and then infer the distribution of diffusion coefficients. One such decay function which is often used is the Kohlrausch-Williams-Watts (KWW) function (otherwise known as a stretched exponential) which is described by [8]

\[ E(b) = e^{-(bD^*)^\alpha} \quad (3) \]

It may be shown that the implied distribution of diffusion coefficients is a Lévy skew alpha-stable distribution [11], [12]:

\[
P(D) = \lim_{\alpha \to 0} \frac{1}{2\pi i} \int_{e^{-i\infty}}^{e^{+i\infty}} e^{-(bD^*)^\alpha} e^{bD} db = \frac{1}{\pi D^*} \int_0^\infty \exp \left\{ -\frac{Du}{D^*} - u^{\alpha} \cos (\pi \alpha) \right\} \times \sin \{u^{\alpha} \sin (\pi \alpha)\} du
\]

A plot of these probability distributions is shown in figure 3 for different values of \( \alpha \). It is clear that the distributions have some degree of ‘normality’ to them with no values of \( D \) being less than 0 as previously desired. Also, as \( \alpha \) decreases it is seen that the ratio of higher to lower diffusion coefficients increases indicating different degrees of negative skewness and kurtosis in the distribution. However, whilst these distribution seem attractive at first glance, they must only be treated as approximations as issues with them are still evident such as the fact that they cannot have more than one peak which may well be expected in a bi-exponential model, they do not allow for any positive skewness and in all distributions (especially those with low \( \alpha \)) extremely high diffusion coefficients approaching infinity are possible. In general the inverse Laplace transform is an extremely ill-conditioned mathematical problem where a minor change in the attenuation curve can have a considerable effect on the implied distribution. The use of a stretched exponential is therefore likely to be due to its convenient mathematical representation and easier fitting to decay curves rather than due to the precision of the implied distributions.

Another approach for analysis of tissue structure using diffusion imaging is known as q-space imaging. This technique uses the assumption that the length of the diffusion sensitising gradients is small \( (\delta < < \Delta) \) so that equation (2) may be re-written in terms of the average probability of a spin moving a distance \( R \) in time \( \Delta \), \( P(R, \Delta) \) [5];

\[
\left. \frac{|S(\tau)|}{S(\tau)|_{g=0}} \right| = \int_{-\infty}^\infty P(R, \Delta)e^{i\gamma\delta g R}dR
\]

If the parameter \( q \) is now defined such that \( q = \gamma\delta g / 2\pi \) (m\(^{-1}\)) then it is clear that the signal attenuation of the signal as a function of \( q \) is the Fourier transform of the average distance probability;

\[
E(q, \Delta) = \int_{-\infty}^\infty P(R, \Delta)e^{i2\pi q R}dR
\]

By obtaining a series of images using different \( q \) values using different direction of \( g \) it is therefore possible to obtain a map of the average propagator \( P(R, \Delta) \) in which effects due to the restriction of diffusion can be seen. However, in order to obtain good accuracy using this technique it is necessary to acquire data using many values of \( q \) over a wide range which will be extremely time consuming and very challenging for the scanner due to the large diffusion gradients involved.

Finally, it should be mentioned that in some cases it is necessary to consider the anisotropy of the diffusion for the tissue under investigation. As previously explained, diffusion anisotropy is neatly summarized using the 3x3, diagonally symmetric...
diffusion tensor, $D$, rather than the scalar diffusion coefficient, $b$. Signal attenuation in Diffusion Tensor Imaging (DTI) is related to the b-value matrix, $b$, by [9]:

$$E(b, D) = e^{-b \cdot D}, \quad b \cdot D = \sum_{i,j} b_{ij} D_{ij}$$

where off-diagonal elements in $b$ may be achieved using diffusion gradients in more than one direction simultaneously [5]. In order to obtain an estimate of the tensor it is necessary to obtain diffusion weighted images using diffusion sensitising in at least 6 directions, although using a greater number of directions will increase the accuracy of the estimation, which may be used to calculate each of the six unknown elements in $D$. Representation of the tensor is a difficult task although some groups choose to use ‘diffusion ellipsoids’ which show the orientation and length of fibers in a 3 dimensional view, as is commonly done in MR tractography of white matter in the brain. However, these processes are somewhat laborious computationally and so two orientation invariant parameters have been reported that summarize both the average diffusivity, regardless of direction, and the level of anisotropy. It can be shown that the trace of a matrix, $\text{Tr}(D)$, is equal to the sum of the eigenvalues for that matrix. As the eigenvalues for $D$ are independent of the orientation of the reference frame in which it is measured, it is simple to see that the trace is also independent. The mean diffusivity is then defined as $\text{Tr}(D)/3$ [9]. However, the use of DTI is limited to tissues which are highly ordered such as the axons in the white matter of the brain. When applying this technique to study diffusion in tumour tissues therefore, the only viable application has been in study of brain tumours [10] as the cellular environment of lesions tends to be very disordered.

III Noise Reduction for Lévy Distributed Processes

Let as assume that the noise generated in an MR image $I$ is the result if a Lévy distributed process. The image can then be taken to conform to the fractional diffusion equation

$$\frac{\partial}{\partial t} I(x, y, t) = D(x, y) \nabla^\alpha I(x, y, t)$$

(1) that includes the Laplacian $\nabla^2$, i.e.

$$\frac{\partial}{\partial t} I(x, y, t) = D(x, y) \nabla^2 [\nabla^{\alpha-2} I(x, y, t)]$$

where

$$\nabla^{\alpha-2} I(x, y, t) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left( \tilde{I}(k_x, k_y, t) \right) \left( \frac{1}{(k_x^2 + k_y^2)^{(2-\alpha)/2}} \right) \exp(ik_x x) \exp(ik_y y) dk_x dk_y$$

and

$$\tilde{I}(k_x, k_y, t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(x, y, t) \exp(-ik_x x) \exp(-ik_y y) dx dy$$

In practice (i.e. for numerical computations operating on a digital image $I_{ij}$), $\nabla^{\alpha-2} I$ is computed using a Discrete Fourier Transform to output the digital equivalent operation which we denote as $\nabla^{\alpha-2}_{ij} I_{ij}$. Forward differencing in time, we can then consider the following iterative filter: For $k = 0, 1, 2, ..., N$

$$I_{ij}^{k+1} = I_{ij}^k + \Delta D_{ij} \left( \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & -4 & 1 \end{array} \right) \nabla^{\alpha-2}_{ij} I_{ij}^k$$

(2) where $I_{ij}^0$ is the input image. Figure 4 show the outputs of the algorithm provided in after 5, 10 and 15 iterations using a time step $\Delta = 0.05$ and a Lévy index $\alpha = 1.98$. The Diffusivity $D_{ij}$ is obtained be applying an edge detector to the image $I_{ij}^0$ to obtain an output $E_{ij}$, say, and compute

$$D_{ij} = 1 - E_{ij}$$

where it is noted that $D_{ij} \geq 0 \forall (i,j)$ and $E_{ij} \geq 0 \forall (i,j)$. In the example given, a Prewitt edge detector has been used [13] The value of $\alpha$ that is used in this case is critical, and must, in general, be close to 2 as given in the example shown in Figure 4. If $\alpha$ moves too far below 2, the lowpass filter $|k|^{-(2-\alpha)/2}$ attenuates the high frequency components in the image too severely at each iteration. On the other had, for values of $\alpha$ close to 2, the number of iterations required to de-noise the image is significantly less than in the application of the non-fractional anisotropic diffusion algorithm, noise reduction being optimal after only 5 iteration as shown in Figure 4.

IV Summary

The principle aim of this has been to describe the nature of diffusion imaging in MR imaging and discuss its use in assessing response of metastatic disease to therapy. It has been shown that the thermal Brownian motion of water may be measured
in vivo using specialized pulse sequences. These sequences cause attenuation in MR images that is dependent on the scanning parameters and diffusivity of the target tissue affording a powerful contrast tool that probes length scales much smaller than those available with other conventional imaging modalities. In most diffusion sequences scanning parameters may be combined into a single 'b-value' which simplifies the equations of signal dependence. By obtaining images with different b-values it is possible to obtain an estimate of the Diffusion Coefficient which characterizes the rate of diffusion in the tissue. In the context of diffusion imaging and given that the implied distribution of diffusion coefficients are Lévy distributed, we have considered extending the method of anisotropic diffusion for noise reduction in digital images to the fractional anisotropic diffusion case. The algorithm considered is compounded in equation (2). This result assumes that the noise generated in a MR image is non-Gaussian and that as $\alpha \to 2$ noise generation becomes the ‘product’ of classical diffusion conforming to Gaussian processes. In this sense, fractional diffusion is a generalisation of classical diffusion which includes long tail distributions associated with Lévy processes associated with the diffusion coefficients studied in MR imaging.

References


