Registration and Statistical Analysis of PET Images Using the Wavelet Transform

Michael Unser[1], Philippe Thévenaz[1], Chulhee Lee[1], and Urs E. Ruttimann[2]
[1]Biomedical Engineering and Instrumentation Program, National Center for Research Resources, National Institutes of Health
[2]Laboratory of Clinical Studies, National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health

Neuroimaging by positron emission tomography (PET) is a major research tool to investigate functional brain activity in vivo. It is a non-invasive method that offers a unique way of measuring the spatial distribution of certain function-specific radiotracers that are injected into the bloodstream prior to imaging. A typical application is to measure cerebral glucose utilization with the tracer [^{18}F]2-fluoro-2-deoxy-D-glucose (FDG), a positron-emitting analog of glucose. Other examples of PET use are the determination of regional cerebral blood flow, and receptor imaging using an appropriate radioactively-labeled ligand [1]. The main difficulty with the analysis of PET data is the absence of reliable anatomical markers and the fact that the shape and size of the brain may vary substantially from one subject to another. The task is further complicated by a low signal-to-noise ratio, and a limited spatial resolution.

In order to gain insight into functional inter-relationships among the various regions of the brain, a variety of statistical methods have been developed. In earlier studies, the data were summarized into a series of regional rates that were typically calculated by integrating the functional data in a number of regions of interest [2]. Although this approach solves the problem of anatomical variability across patients, it is extremely time-consuming because each individual region needs to be specified manually, requiring the involvement of a highly skilled operator, without any guaranty of reproducibility. Furthermore, investigators are often tempted to define more regions than can be analyzed with sufficient statistical power, with the consequence that most reports do not correct the significance levels for multiple testing.

A more recent trend has been to look at more global solutions to the problem, based on the premise that the images could all be mapped spatially onto some common reference, such as the standardized stereotaxic coordinate framework of Talairach, et al. [3]. Once the brain images are aligned with respect to each other, it is possible to develop image-based statistical methods for detecting differences between subject groups. One of these approaches is the statistical parametric mapping (SPM) method of Friston, et al., which achieves registration through spatial warping [4, 5]. This type of approach is faced with two major difficulties that are somewhat interrelated. The first is the problem of registration [6, 7]. Is it entirely reasonable to assume that there exists a geometric transformation that maps a particular brain onto a common standard? A proper deformation model is not presently known. Using a model that has too many degrees of freedom may strongly bias the data towards the reference and obliterate meaningful differences to be detected. On the other hand, without sufficient model flexibility, the different brain regions may fail to be properly aligned. The second difficulty with the global approach consists in the development of analysis methods that have a sound statistical foundation. This may not be an easy task because pixels in the spatial domain are statistically not independent, a problem that is further aggravated by the nature of the preprocessing transformations applied.

The purpose of this article is to present an overview of our own methodology. The two important processing steps are 1) the registration of the individual brain images, and 2) the subsequent statistical analysis in order to detect differences in functional activity between subject groups. What is remarkable is that the wavelet transform plays an important role in each of them, but for different reasons. The registration algorithm that is described below uses the idea of a multiresolution analysis, which consists of representing a signal by a sequence of fine-to-coarse continuous functions, providing approximations at various resolutions [8]. Our approach is to match images

0739-5175/95/$4.00©1995

onto a common standard by global affine transformation, using a multiscale extension of the Levenberg-Marquardt non-linear optimization method. For the registration problem, it is the concept of an underlying continuous signal model that is especially important. Polynomial spline image representations are particularly attractive in this context because of their many useful properties [9]. The statistical analysis method that is presented below uses the wavelet transform because of its unique space-frequency localization properties. This last application emphasizes the discrete interpretation of this transform and relies on the orthogonality property. The main advantage is that the discriminative information gets concentrated into a relatively small number of wavelet coefficients, which can be identified through statistical testing. The image reconstructions obtained from this reduced coefficient set provide a relatively uniform and noise-free visualization of the significant between-group functional differences.

Overview of the Procedure

Before discussing the individual components of the system, we provide a brief overview of the experimental procedure.

Subjects and Image Acquisition

The PET images used for this study represent cerebral glucose utilization rates as determined using the radioactive tracer FDG, and were obtained from 10 patients with alcoholic organic mental disorders and 8 normal volunteers. For each subject, 21 slices with 128x128 pixels (pixel size = 2mm) were acquired with the NeuroPET scanner, providing transverse and axial spatial resolution of 7 and 11.5 mm full width at half maximum (FWHM) (of the point spread function), respectively, with an interslice separation of 4 mm.

Image Registration

The first step was to register these volumetric PET data sets with respect to each other. In our case, we considered single characteristic slices that clinicians had pre-selected for the analysis (2-D processing). These images were then matched to a common standard at each slice level by using the multiresolution registration procedure described below. The class of geometrical transformations was restricted to the family of affine transformations; i.e., any combination of translation, rotation and anisotropic scaling. The same approach can also be extended for performing a full 3-D registration.

Statistical Analysis

Each group is then characterized by its mean image, which is obtained by averaging. Subtraction of the mean-images yields the information of clinical interest, that is, the depiction of functional differences between the alcoholic patients and the normal volunteers. Additional relevant statistical information is provided by the within-group variance estimates. Testing for significant differences in the image domain is difficult because of the amount of noise and the large number of variables (128x128) that need to be considered. Instead, we chose to perform the statistical analysis in the wavelet transform domain, as described below. The main advantage of using the wavelet transform is that it concentrates the discriminative information into a relatively small number of coefficients, without a significant loss of spatial resolution. It is thus possible to reduce the number of variables that need to be tested, while also taking advantage of the corresponding increase in signal-to-noise ratio.

Multiresolution Image Registration

The registration problem is best stated in a continuous framework. Let $s(\boldsymbol{x})$ and $r(\boldsymbol{x})$ denote the object and reference images, respectively, which are both represented as functions of the continuous spatial variable $\boldsymbol{x} = (x_1, x_2)$. Although we limit our presentation to the two-dimensional case, we have also implemented and applied the procedures to the three-dimensional case. The purpose of registration is to find a spatial mapping, T, that geometrically transforms s into an image, $s_T(\boldsymbol{x}) = (\boldsymbol{T}_T s)(\boldsymbol{x})$, that is maximally compatible with the reference, r. Assuming that we can define a meaningful metric, $d(s_T,r) = \|(\boldsymbol{T}_T s)(x) - r(x)\|$, that measures the similarity between s_T and r, the task is then to determine the transformation T_T that minimizes $d(Ts,r)$ among a certain class of admissible transformations T. It turns out that even in the simplest case of translational and rotational alignment, this optimization problem does not have a simple solution; it usually necessitates the development of iterative search techniques, which can be computationally quite expensive [6, 10, 11]. In addition, such simple rigid body transformations are only rigorously valid for aligning PET images from the same subject.

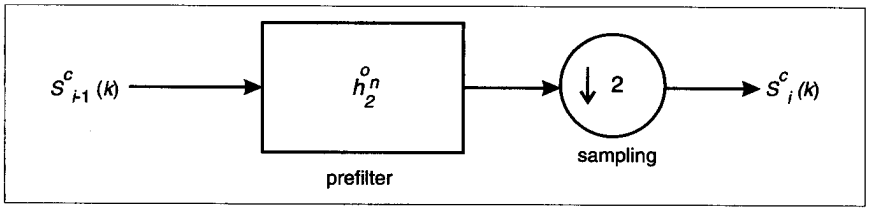
Here, we have chosen to consider the extended class of affine transformations that can also account for more general distortions, such as shear or change in aspect ratio. The general transformation model is given by:

$$(Ts)(\boldsymbol{x}) = s(\boldsymbol{T}(\boldsymbol{x} - \boldsymbol{t})) \qquad (1)$$

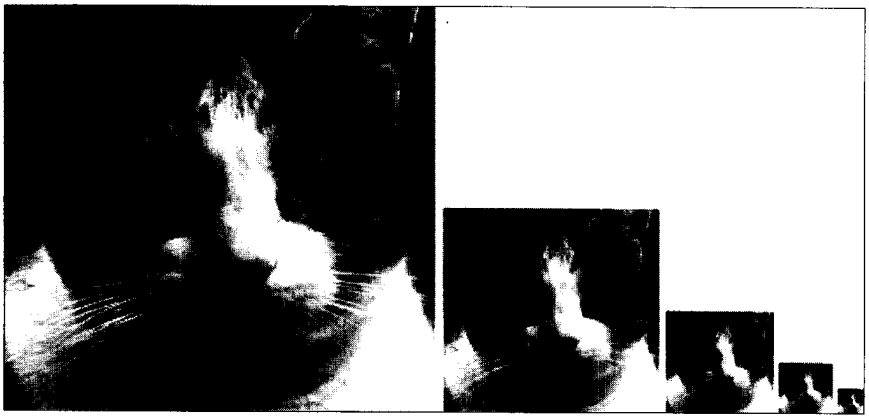
where $\boldsymbol{T}$ is a 2x2 (3x3 in 3-D) non-singular matrix and $\boldsymbol{t}$ a translation vector. We selected this particular type of deformation because it has a reasonably small numbers of parameters (4+2 parameters in 2-D, and 9+3 parameters in 3-D), and yet it appears to offer enough flexibility for performing the most important anatomical corrections. In addition, it results in a continuous and well-behaved mapping between s and r that has a simple physical interpretation. We feel that having a rather constrained deformation model is crucial to this application because we do not want to over-distort our images nor introduce artifacts that were initially not present in the data. In other words, we want to stay as close as possible to the data. The adequacy of the model can always be checked, a posteriori, by examining the various group averages as well as the error images. The only risk with this approach is that it may smooth out certain characteristic patterns that occur in regions of the brain that are not well aligned across patients.

The Polynomial Spline Pyramid

The registration problem, as it has just been formulated, assumes that the images s and r are continuously defined. A convenient way of obtaining such representations is to fit our initial pixel arrays with polynomial spline surfaces [12]. Such spline surfaces are made up of rectangular, piecewise, polynomial sections of degree n that are connected together in a way that guarantees the continuity of the function and its derivatives up to order n-1. The main reasons for selecting such a spline model are the following: First, the model is uniquely determined by the value of the surface at the grid points (cardinal representation), essentially preserving the 2-D structure of the data. Second, polynomial splines have a simple explicit form that makes them easy to manipulate; in particular, it is straightforward to re-sample the model (geometric transformation) or to compute its gradient at the grid points, as will be required by our optimization procedure. Finally, they are ideally suited for multi-scale processing because of their multiresolution properties [9].



1. Reduction procedure for the spline pyramid.



2. Example of cubic spline image pyramid (PET/CAT scan).

At the finest resolution level, $i = 0$, where the step size is one, the polynomial spline representation is exact in the sense that it provides a perfect fit. The corresponding model parameters are the initial pixel values,

$$s^c_o(k,l) = s(x,y)|_{x=k,y=l}$$

The polynomial spline pyramid corresponds to the sequence of fine-to-coarse signal approximations $\{s_0 = s,...s_i = A_i s,...s_I = A_I s\}$, where A_i represents the projection (or approximation) operator at resolution level i, and I is the depth of the pyramid. Each of the functions $s_i(x)$ corresponds to a spline approximation with a step size 2^i and is entirely characterized by its samples values at the grid points

$$s^c_i(k) = s_i(x)|_{x=2^i k}$$

These coefficients are computed iteratively by digital filtering and decimation, as illustrated by the block diagram in Fig. 1. The least squares prefilter, h^n_2 acts as an anti-aliasing filter; as the order of the spline n increases, it converges to the ideal half-band lowpass digital filter [13, 14]. As an example, the pyramid representation of the test image PET/CAT scan is shown in Fig. 2.

Higher order splines usually result in better quality approximations, although they require more computations and give rise to more complicated formulas. For our experiments, we used a cubic spline representation, which provides a good cost/performance compromise. Figure 3 represents an enlarged representation of the coarser level of the pyramid for both piecewise constant and cubic spline models. Note how the quality of the approximation improves with the higher order model, which provides a smoother surface representation. The zero model is not appropriate here because it is not even differentiable.

Multiresolution Alignment Procedure

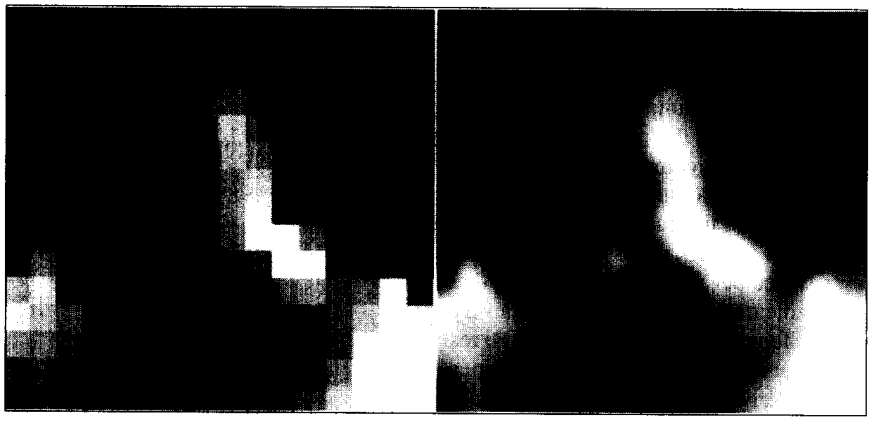
Our measure of similarity between images is the L_2-norm of the difference. Thus, the goal is to find the transformation T that minimizes the integrated square error between r and Ts. Our approach is to cast this optimization problem into a multiresolution framework and to consider the signal and reference approximations, $s_i = A_i s$ and $r_i = A_i r$, in a spline pyramid with levels $i=0,...,I$. Starting at the coarsest level, I, we successively solve the sequence of problems for $i=I$ down to 0:

$$(T_i, t_i) = \arg\min \|Ts_i - r_i\|^2 \qquad (2)$$

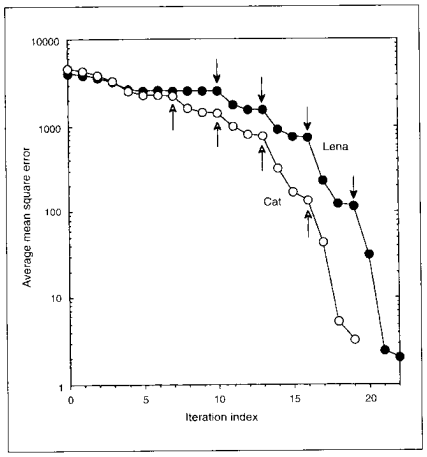
In fact, we use a discretized version of this error obtained by sampling these functions at the grid points. The optimum alignment parameters at scale i are determined iteratively by considering a perturbation $(\Delta T, \Delta t)$ on the previous solution. This procedure ultimately yields the solution to our initial problem when $i=0$. The approximation problem at a given scale is solved by using a variation of the Marquardt-Levenberg least-squares optimization algorithm, which is iterative and requires the evaluation of the partial derivatives of s_i with respect to ΔT, and Δt. This algorithm first mimics a standard steepest descent (or gradient) procedure, and progressively switches to an inverse-Hessian method as the minimum is approached [15]. An important algorithmic simplification is that the optimization is always performed around the operating point $(T, t) = (I, 0)$, using the fact that the error criterion can also be rewritten as:

$$\varepsilon^2 = \|Ts_i - r_i\|^2 = \frac{1}{|T|}\|s_i - T^{-1} r_i\|^2 \qquad (3)$$

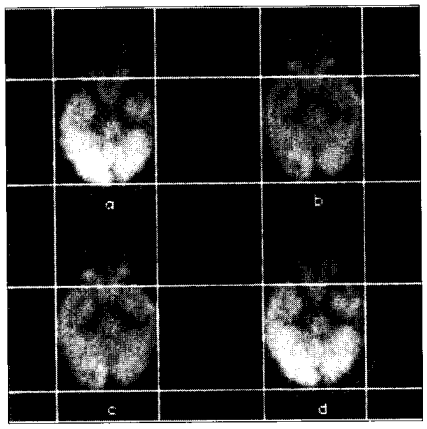
where $|T|$ denotes the determinant of the matrix T. In other words, we can account for the current transformation estimate by modifying the reference appropriately as we go along with the optimization. The



3. Comparison of zero order and cubic spline models at the coarser level of the pyramid (x16). (a) enlarged zero order approximation, (b) enlarged cubic spline approximation.



4. Evolution of the quadratic error per pixel as a function of the iteration number during the registration of the two test images, PET/CAT scan and Lena.



5. Example of registration of a 128x128 PET image. (a) unregistered data, (b) reference, (c) reference, (d) registered data.

only difficulty with this approach consists in the updating of the current transformation, T, which requires the use of special pseudo-commutativity rules. At each iteration, the modified reference $T^{-1}r_i$ is evaluated by re-sampling the underlying spline surface at the new location of the grid points; this computation uses the B-spline representation of this function. The advantage of this modified scheme is that the Hessian matrix needs only to be computed once per resolution level, and that the gradient of the error criterion is a very simple function of the spatial derivatives of $s_i(x)$ at the grid points. These quantities can be derived by simple digital filtering [16]. This makes the method substantially faster than the standard Marquardt-Levenberg scheme. A detailed account of the algorithm for the special case of a rigid transformation (rotation + translation) can be found in [17]; the more general affine case will be described in more detail in a forthcoming article.

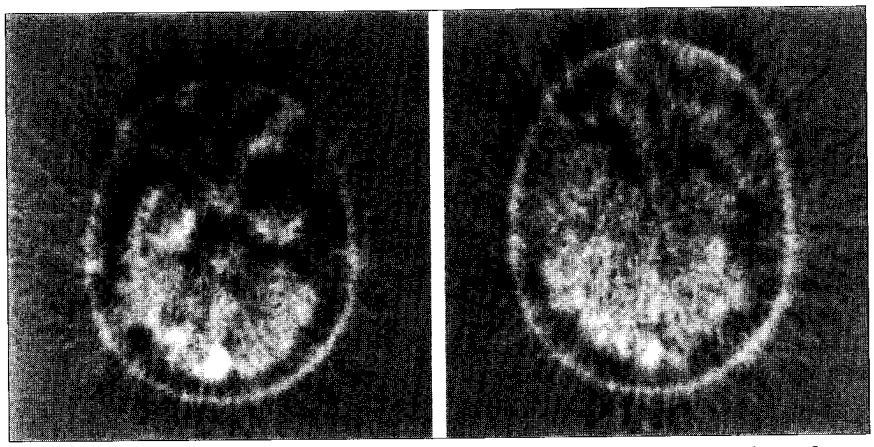
Registration Results

We first checked the performance of the algorithm by applying it to some test images that were pre-distorted by translating, rotating, and scaling a standard reference using cubic spline interpolation. In particular, we monitored the evolution of the normalized quadratic error criterion, Eq. 3, as a function of the iteration step. Two examples of such curves are shown in Fig. 4. The transformation parameters for both images, ("PET/CAT" scan and "Lena") were $\boldsymbol{t}$ =(5,10), θ=15° (rotation angle) and λ=0.8 (isotropic reduction factor); the processing was performed using a five-level pyramid ($I = 4$). The translation factors were initially set to zero, and the matrix $\boldsymbol{T}$ set to identity. The arrows in Fig. 4 indicate the places of transition to the next level. It can be seen that most iterations are spent at the coarsest level of the pyramid and that the initial improvement is quite progressive. Once the algorithm switches to the next finer scale, the convergence is extremely fast. This near-optimal behavior is maintained until the end. Note that a minimum of two iterations per scale was imposed in order to check convergence.

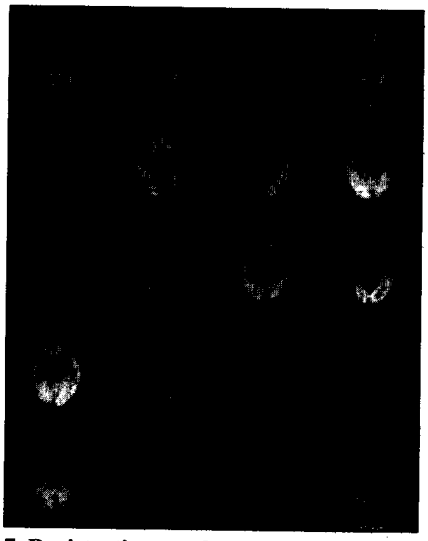
In all the noise-free examples that we considered, the parameters were always estimated with almost perfect accuracy: $\triangle t_x, \triangle t_y < 10^{-3}$ pixels, $\triangle\theta < 10^{-3}$ degrees, and $\triangle\lambda < 10^{-4}$. For comparison, we also tried a single scale optimization, but this procedure typically did not converge unless the initial guess was within two pixels of the true solution. Even in this very favorable case, the single-scale approach took much longer than the multiresolution implementation.

We also investigated the behavior of the algorithm in the presence of Gaussian additive noise [17]. The procedure turned out to be extremely robust and its convergence properties were essentially unaffected. As expected, we observed a slight loss of accuracy in the parameter estimates as the amount of noise was increased. However, even under the most adverse conditions (SNR = 0 dB), the quality of the estimation remained remarkably good, yielding image registration within a few tenths of a pixel.

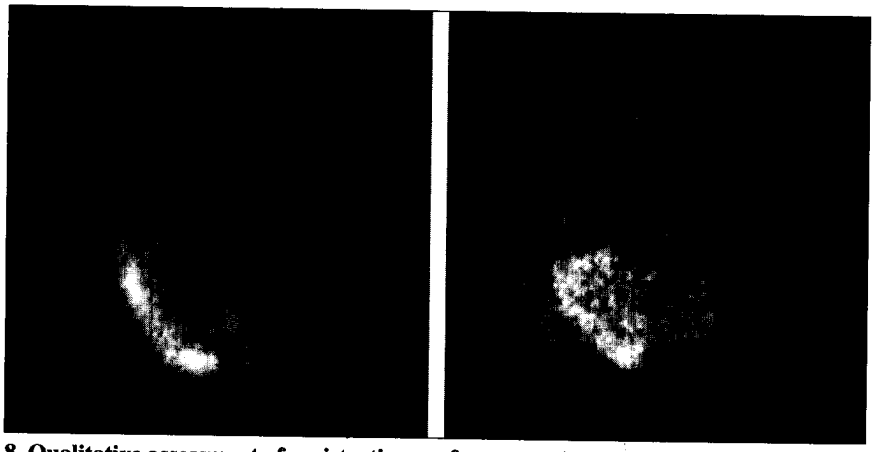
Next, we considered the application of the procedure to real PET data. Figure 5 shows one instance of such a registration. Here, the effect of the transformation has been to stretch the original data along the vertical axis, and somewhat along the horizontal as well, as can be seen by comparing section (1) with section (4). The reference is shown twice in order to facilitate the visual comparison, and a grid has been added. The estimated transformation matrix is $\boldsymbol{T}$ = ((1.053, 0.014), (-0.023, 1.093)), with a translation given by $\boldsymbol{t}$ = (1.7, -6.6). Figure 6 displays the corresponding difference images. Note how the error associated with the size mismatch between data and reference disappears in the registration process, while the overall quadratic error is reduced by a factor of approximately two. Figure 7 shows the registration results for a whole data set, together with the overall group average. The stripe patterns and the oval frames over some of the images are artifacts generated by the acquisition and reconstruction process. To suppress the influence of



6. Difference images corresponding to Fig 5. (a) difference between unregistered data and reference, (b) difference between registered data and reference.



7. Registration results for the full data set including 10 alcoholics subjects (labeled 2, 4, 6, 8, 11, 12, 14, 15, 16, 17), and 8 normals (labeled 1, 3, 5, 7, 9, 10, 13, 18). The average of the whole data set is shown last.



8. Qualitative assessment of registration performance. (a) subtraction of the group-averages (unregistered case), (b) subtraction of the group averages (registered case).

extracranial pixels, we first computed a mask by thresholding a strongly lowpass filtered version of the image. The optimization was then performed selectively over the masked region of interest. Figure 8 displays the average-differences between the alcoholic and normal groups for the unregistered and registered case. This final result indicates that the registration has been truly beneficial. For example, looking at the upper part of the image (the frontal lobe), one can see that many features associated with potentially imperfect registration have vanished. At the same time, an area of enhanced activity in alcoholic patients, at the back of the brain (cerebellum), shows up more homogeneously. Furthermore, the highly contrasted spot at the very back of the brain (cerebellar vermis) stands out much more clearly in the registered image than in the unregistered one.

Discussion

Based on our experimental results, it appears that the algorithm performs extremely well, at least in finding the best affine mapping between an image and a reference. The main computational features that should be emphasized are the following.

The algorithm provides extremely accurate estimates of the registration parameters. In a noise-free environment, the alignment is essentially perfect. This is a consequence of a consistent design that uses a continuous polynomial spline image model.

The multiresolution algorithm is superior to its single-scale version in a number of respects. First, it converges much more rapidly because the spatial resolution of the underlying image model is adapted to the step size of the algorithm. Second, it is much less likely to get trapped in a local minimum because of the smoothing effect of the pyramid. Finally, it is much faster because most iterations are performed at the coarsest level of the pyramid.

The method is iterative and is therefore somewhat sensitive to the quality of the initial guess. However, we found that the algorithm could easily handle angular discrepancies as large as 20 degrees, and translational errors of the order of 2^I pixels, where I is the depth of the pyramid.

The algorithm is very robust to measurement noise. Even when the signal-to-noise ratio is as low as 0 dB, it is still possible to obtain parameter estimates within a few tenths of a pixel, and less than one tenth of a degree. Noise has almost no effect on convergence because of the noise reduction properties of spline pyramids.

While the multiscale registration algorithm performs to satisfaction, it is still somewhat unclear whether or not the class of transformations that has been considered is sufficiently powerful to accommodate the most important morphological differences between subjects. The variability of the images available for our study is rather large, as can be seen from Figs. 5-8. These discrepancies have different origins, ranging from head positioning mismatches between the patients, imperfect calibration of the intensity scale, detector dropouts during acquisition, unexpected or uncontrolled stimulation during the scans, different brain geometry between patients, different individual activity patterns, or different activity patterns correlated with the group assignment (alcoholic or non-alcoholic), to cite a few. The purpose of our registration procedure is to correct for some of the geometric factors as they influence the representation within the image plane. However, out-of-plane tilt, for example, was not corrected for by these strictly 2-D procedures. Performing a true 3-D registration should overcome some of these limitations. There is also the possibility of extending the class of allowable transformation by introducing additional parameters. Presently, while we are still in the process of acquiring experience with real data sets, we prefer to work with a more conservative, yet controllable class of deformation models. The risk is that we may miss certain functional differences, but we also reduce the likelihood of generating false positives.

Statistical Wavelet Analysis

Once the images have been registered and the group-averages computed, we take the wavelet transform of the between-group differences and perform the statistical analysis in the transformed domain, as suggested in [18]. In the next section, we provide a brief review of the wavelet transform, emphasizing the properties that are relevant for this particular application.

The testing procedure is described below with some examples of PET data analysis provided.

Wavelet Transform

The wavelet transform decomposes a signal into a set of orthogonal components describing the signal variations across scales [8, 19]. Specifically, a signal, s, is represented by its wavelet expansion (synthesis formula) as:

$$s = \sum_{k\in Z} c_I(k)\phi_{I,k} + \sum_{i=1}^{I}\sum_{k\in Z} d_i(k)\psi_{i,k} \quad (4)$$

where $\theta_{I,k}$ and $\psi_{i,k}$ are the orthogonal scaling and wavelet basis functions. The expansion coefficients $c_I(k)$ and $d_i(k)$ in Eq. 4 are obtained by inner product (analysis formulas:

$$c_i(k) = < s, \phi_{I,k} >, \quad d_i(k) = < s, \psi_{i,k} > \quad (5)$$

The wavelet basis functions are typically constructed by dilation (scale index i) and translation (index k) of the wavelet function $\psi(x)$.

$$\psi_{i,k} = 2^{-i/2}\psi(x/2^i - k) \quad (6)$$

The first term in Eq. 4 provides a coarse approximation of our signal at the scale I, which corresponds to the lowest level of the pyramid in Fig. 2. The wavelet term represents the residual error between two approximations at scales i-1 and i; that is, the difference between two successive levels of the pyramid (cf. Fig. 2). Although the synthesis and expansion formulas, Eqs. 4 and 5, are usually given for continuous signals [8, 20], equivalent expressions also exist for a purely discrete framework [21]. In the discrete formulation, which is the appropriate one here, these formulas can be rewritten in the following matrix form:

$$s = \boldsymbol{W}^T \boldsymbol{d} \quad (7)$$

$$\boldsymbol{d} = \boldsymbol{W}s \quad (8)$$

where $\mathbf{s} = (...,s(k),...)$ is the (infinite dimensional) signal (or image) vector, $\boldsymbol{W}$ the orthogonal wavelet transformation matrix, and $d = (..,d_I(k)_I...d_I(k)_I...i_I(k)_I...)$ the wavelet coefficient vector. The wavelet transform (Eq. 8) is therefore an orthonormal transformation of our image array s.

Rather than defining the transform matrix $\boldsymbol{W}$ explicitly, it is much easier to describe the underlying decomposition algorithm, which uses two complementary filters h and g. The lowpass filter h satisfies the so-called quadrature mirror filter (QMF) conditions:

$$|H(\omega)|^2 + |H(\omega+\pi)|^2 = 1 \quad (9)$$

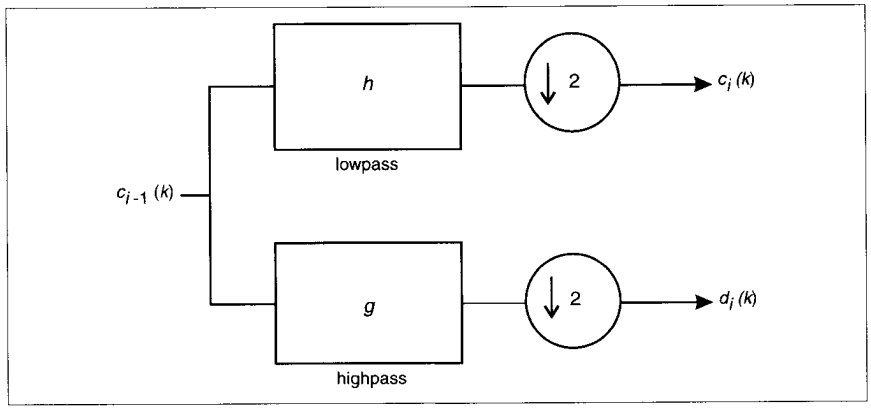
$$H(0) = 1 \Rightarrow H(\pi) = 0 \quad (10)$$

where $H(\omega)$ is the Fourier transform of h. The highpass filter g is the modulated version of h given by

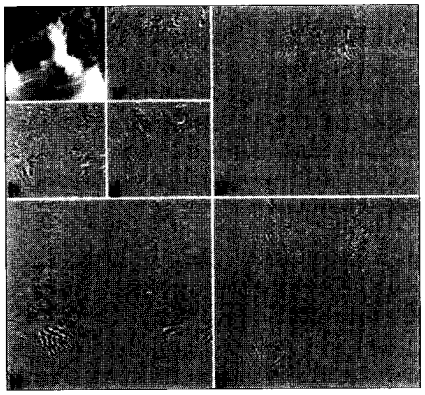
$$g(k) = (-1)^k h(1-k) \quad (11)$$

The wavelet decomposition is implemented iteratively by successive filtering and decimation using the QMF filterbank described in Fig. 9 [8]. Note the similarity between the lowpass branch of the algorithm (filter h) and the pyramidal algorithm in Fig. 1. The decomposition is easily extended to 2D through the use of tensor product basis functions, which essentially amounts to applying the 1D decomposition algorithm successively along the rows and columns of the image. An example of wavelet transform with a depth of I = 2 of our test image is shown in Fig. 10; each wavelet sub-image (or channel) had its gray scale linearly expanded for maximum contrast display. This example illustrates the property that the wavelet channels in the upper right (V), lower left (H), and lower right (D) quadrants tend to amplify high resolution (or, equivalently, high frequency) vertical, horizontal and diagonal edges, respectively. The same qualitative behavior also applies for the mid-range frequency components (v, h, and d). The lower resolution image in the upper left corner is equivalent to the level 2 image in the spline pyramid in Fig. 2. The information at the various scales corresponds to different frequency bands, but the wavelet coefficients retain a certain degree of spatial localization.

It is the space/frequency localization property of the basis functions that make the wavelet transform especially attractive for this particular application. It is reasonable to expect that the group-differences in functional activity will be rather localized in space, but with slow changes in intensity (essentially lowpass). For these reasons, most of the signal contribution will be concentrated in a relatively small number of wavelet coefficients. The noise, on the other hand, will spread out more evenly throughout the wavelet space. White noise, in particular, will remain white within the different resolution channels because the transformation was selected to be orthogonal. It is thus possible to discard most of the non-significant coefficients. This process is somewhat analogous to applying a lowpass filter to the inter-group difference image, with the essential difference that the lowpass cutoff frequency does not have to be specified beforehand, but is rather determined by statistical testing. This also achieves a useful data reduction, while essentially preserving the correlation structure of the data.



9. Fast wavelet transform decomposition algorithm.



10. Example of two-dimensional wavelet transform (PET/CAT scan).

When the spectral density of the signal is non-increasing, the optimal wavelet transform for maximum energy compaction turns out be the ideal bandpass decomposition (modulated-sinc wavelet) [22]. Accordingly, we should expect wavelet transforms with good bandpass characteristics (i.e., larger regularity indexes or higher order splines) to be the most efficient at detecting significant metabolic differences. In practice, the achievable energy compaction has to be counterbalanced by the fact that higher order wavelets tend to be less localized and that bandlimited signal approximations also give rise to Gibbs oscillations. The issue of selecting the best transformation is therefore essentially a matter of compromise.

Another important consideration for the efficiency of the statistical procedures to be presented below is the correlation structure of the wavelet coefficients. For stationary noise, the correlation between two wavelet channels is a function of their spectral overlap, and is zero for the ideal bandpass decomposition. In the non-ideal case, the correlation can usually be assumed to be negligible since it decays as $\Delta^{n+3/2}$, where Δ is the scale ratio between resolution levels and n the number of vanishing moments of the transform [18].

For our experiments, we selected orthogonal spline (or Battle-Lemarié) wavelets [8, 23, 24]. This family of transforms has the advantage of using symmetrical basis functions; it also provides a simple way of reducing spectral overlap by increasing the degree of the spline n. In the limit as n goes to infinity, the Battle-Lemarié wavelets tend to the ideal bandpass filter [24, 25]. These transforms were computed using the fast wavelet transform algorithm shown in Fig. 9. Explicit filter formulas and a discussion of important implementation issues such as the specification of boundary conditions can be found in [26].

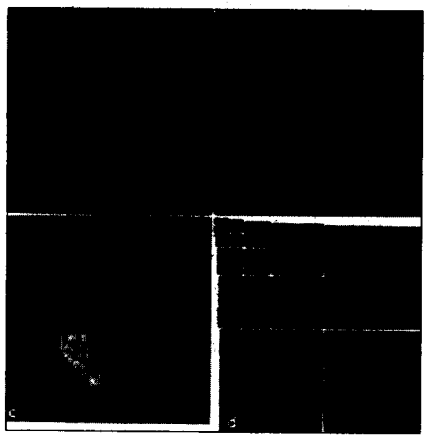
Statistical Analysis

The relevant clinical question is whether or not there are statistically significant differences between the two group-average images, or, equivalently, whether the test image represented by the group-average difference is different from noise. Here, for simplicity, we assume that the noise is white with zero mean and standard deviation, σ. The standard deviation was computed by pooling the within-group variances estimated on a pixel by pixel basis. Note that the procedure can be extended for the independent non-stationary case by testing the data against a simulated null image with a local variance that is proportional to the corresponding pixel-by-pixel within-group variance estimate [18].

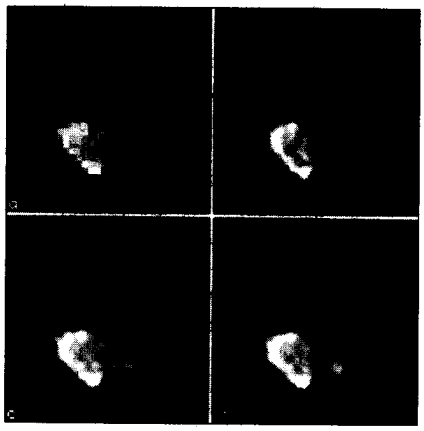
Results for our PET data are summarized in Fig. 11. Images (a) and (b) represent the group-averages for the normal and alcoholic subjects, respectively. Subtraction of those two images ((b)-(a)) yields the difference image, which is shown in Fig. 11c. An equivalent representation of this information is provided by the cubic spline wavelet transform in Fig. 11d.

In the first step of the analysis, we look at the various wavelet channels individually and perform omnibus tests to determine whether or not they represent noise only. These tests are independent because, for all practical purposes, the wavelet channels are not correlated with each other, as indicated previously. Due to the assumption of white noise and the orthogonality of the wavelet transform, the test statistic for each channel is a chi-square with m degrees of freedom, where m is the number of wavelet coefficients in the channel. There are 3 channels at each scale i, each yielding $m_i = (128/2^i)^2$ coefficients. With 5 resolutions to be analyzed, there are 15 tests to be performed. To maintain an experimentwise error rate of 5%, we make the Bonferroni adjustment for multiple testing, and select the significance level for each channel at $\alpha = 0.05/15$. For our particular data set, irrespective of the degree of the spline wavelets, none of the null hypotheses were rejected at the two finer scales and all were rejected below that resolution. Therefore, 3/4+3/16 = 15/16 of the total number of coefficients could be discarded out of hand as encoding image noise only.

Once the data have been found to be significant and reduced to a statistically significant bandwidth, we conduct a finer analysis to locate specific differences in the image. For this purpose, the remaining coefficients are tested each separately using a two-sided z-test at $\alpha = 0.05$, which, in this case, is equivalent to setting a simple threshold at $T = 1.96\,\sigma$. This test statistic follows because the wavelet transform is a linear operation, and the linear transform of a Gaussian variate (the noise) is itself Gaussian distributed. After application of this testing procedure, the total number of wavelet coefficients found to be significantly different from noise were



11. Alcoholics versus normals. (a) group-average for the normal subjects, (b) group-average for the alcoholic subjects, (c) difference image [(b)-(a)], (d) wavelet transform of (c) [$I = 5$ and $n = 3$].



12. Resynthesis of the between-group difference image in Fig. 11c using significant wavelet coefficients only. (a) zero order (Haar), (b) linear spline, (c) cubic spline, (d) quintic spline.

217 (1.32%), 245 (1.49%), 249 (1.52%), and 251 (1.53%), for spline wavelets of order 0, 1, 3, and 5, respectively.

Figure 12 displays the reconstructed images obtained by retaining only the significant coefficients and applying the inverse wavelet transform. The zero-order resynthesis exhibits some characteristic blocking artifacts, while the first-order image contains some linear streaking artifacts. The third-order and fifth-order reconstructions are visually very similar and have a much smoother appearance.

Discussion

The main difficulty with the statistical analysis of PET data sets is the high dimensionality of the feature space (128x128 pixels per image). The wavelet

transform clearly reduces this problem due to the preliminary data reduction that is achieved through omnibus testing. The decorrelation of the resolution channels enables the application of statistical procedures that are more rigorous and more powerful than pixel-based testing in the spatial domain. Statistical power, and thus, greater sensitivity, is obtained because of the partitioning of the data into wavelet channels. Most of the discriminative information tends to be concentrated in the lower resolution channels, which have very few coefficients and a low variance due to the small bandwidth. This reduces both the number of tests and the standard error.

The wavelet transform also offers the possibility of a finer level of analysis for locating the major differences in metabolic activity. In this example, the series of follow-up tests was not protected for total false-positive error rate and the results of wavelet re-synthesis should therefore be treated with some caution. Re-synthesis using only wavelet coefficients retained after Bonferroni-adjusted significance testing (about 0.3% of the coefficients) still resulted in the display of meaningful regions of metabolic rate differences, very similar to those shown in Fig. 12. However, the obvious trade-off in selecting increasingly stricter significance levels is that the estimates of the actual magnitudes of these differences become more and more biased toward lower values. Evaluating the exact statistical power of the proposed procedures and the bias of the estimates of change will require Monte Carlo simulation studies

Based on our experiments, the most appropriate decomposition appears to be the Battle-Lemarié cubic spline wavelet transform. Since increase of the spline order incurs a loss of localization, there is no compelling reason for using higher order splines. As far as the statistical procedures are concerned, we feel that there is still room for improvement. The first extension would be to use a slightly more sophisticated colored noise model, which would still allow us to take advantage of the channel decorrelation property of the wavelet transform. Alternatively, one should also explore non-stationary models which appear to be better suited to characterizing the noise contributions due to misregistration artifacts and counting statistics. The simplest technique that accounts for the within-group variability of the data on a coefficient-by-coefficient basis is a local t-test that could be applied in the wavelet transform domain directly. Another more general approach based on the actually observed data would be to generate empirical distributions for appropriate test statistics of individual wavelet coefficients by using randomization test procedures [27]. Such nonparametric procedures would have the desirable property of fully taking into account the correlation structure of the image data, which could be locally varying and still be decomposed by the wavelet transform. This approach would retain the advantage of concentrating the useful information into a relatively small number of coefficients without much compromise in spatial resolution.

Conclusion

In this article, we have described a general procedure for the processing and analysis of PET data. We have used the multiresolution framework of the wavelet transform to derive new solutions for the two main processing steps.

The first task was to align the various brain images using a general affine deformation model. Our registration procedure uses a continuous polynomial spline image model and takes advantage of the multiresolution structure of the underlying function spaces. This method implements a non-linear least squares optimization technique with a coarse-to-fine iteration strategy that substantially improves the overall performance of the algorithm. Other biomedical applications that could benefit from this algorithm include averaging techniques for noise reduction in high resolution electron-micrographs [11, 28], multimodality imaging [7], and the alignment of autoradiographic slices for three-dimensional volume reconstruction.

The second task was to analyze the series of registered images and to detect the between-group differences in metabolic brain activity. We chose to take advantage of the orthogonality and localization properties of the wavelet transform. Our approach was to apply this transform to the group-difference image and identify the wavelet channels that are globally significantly different from noise. These differences were then localized spatially through a follow-up analysis, which considered the reduced feature set provided by the first analysis, and tested for the coefficients that are statistically significant; these typically represented a very small fraction of the initial set (1.5%). Resynthesis of group-difference images based on the significant wavelet coefficients only resulted in images displaying relatively uniform, noisy-free regions of glucose utilization differences, with very little reconstruction artifacts.

Michael Unser (M'89-SM'94) was born in Zug, Switzerland, on April 9, 1958. He received the M.S. (with highest honors) and Ph.D. degrees in Electrical Engineering in 1981 and 1984, respectively, from the Swiss Federal Institute of Technology in Lausanne, Switzerland.

He is currently a Visiting Scientist with the Biomedical Engineering and Instrumentation Program, National Institutes of Health, Bethesda, which he joined in 1985. He has also been affiliated with INSERM (French National Institutes of Health and Biomedical Research) since April 1988. His research interests include the application of image processing and pattern recognition techniques to various biomedical problems, multiresolution algorithms, wavelet transforms, and the use of splines in signal processing.

Dr. Unser was awarded the Brown-Bowery Price in 1984 for his work on texture analysis and automated inspection. He is the author or co-author of more than 50 jounal papers. He serves as an Associate Editor for the IEEE Transactions on Image Processing, IEEE Signal Processing Letters, and is a member of the Multidimensional Signal Processing Committee of the IEEE Signal Processing Society. He is also on the Editorial boards of Signal Processing and Pattern Recognition. Dr. Unser can be reached at: Biomedical Engineering and Instrumentation Program, Building 13, Room 3N17, National Institutes of Health, Bethesda, MD 20892-5766, USA. Email: unser@helix.nih.gov

Philippe Thévenaz was born in Lausanne, Switzerland, on December 16th, 1962. In 1981, he began his university studies at the Lausanne Swiss Federal Institute of Technology (EPFL), where he graduated with a diploma in microtechnology in January 1986, he then joined the Institute of Microtechnology of the University of

Neuchâtel, Switzerland, where he obtained a PhD degree in June 1993, with a thesis on the use of the linear prediction residue for text-independent speaker recognition, advisor. He then went to the USA where he joined the National Institutes of Health, Bethesda MD, where he is currently involved with biomedical image registration and fast algorithms for geometric image transformations.

Chulhee Lee received the B.S. and M.S. degrees in Electronics Engineering from Seoul National University in 1984 and 1986, respectively, and the Ph.D. degree in Electrical Engineering from Purdue University, West Lafayette, IN, in 1992. From 1986 to 1987, he was a Researcher in the Acoustic Laboratory at Technical University of Denmark (DTH), where he participated in a signal analysis project. He is currently a Visiting Fellow with the Biomedical Engineering and Instrumentation Program, National Institutes of Health, Bethesda, MD. His research interests include image/signal processing, pattern recognition, neural networks, and AI. Dr. Lee is a member of Tau Beta Pi, Eta Kappa Nu and KSEA.

Urs E. Ruttimann received his undergraduate degree in EE from the Eidgenossische Technische Hochschule (ETH), Zurich, Switzerland (1966), and his Ph.D. in biomedical engineering from the University of Pennsylvania (1972). In 1972 he became Assistant Professor in Clinical Engineering at the George Washington University, since 1980 he held academic appointments with the Department of Child Health and Development, and in 1993 became Research Professor of Pediatrics. From 1980 to present he has been a Biomedical Engineer at the National Institutes of Health, first with the National Institute of Dental Research, and since 1991 with the National Institute on Alcohol Abuse and Alcoholism. His research interests include biological signal analysis, 3D medical imaging, and quantitative outcome analysis from intensive care. Dr. Ruttimann is a member of the IEEE Engineering in Medicine and Biology, the Acoustics, Speech, and Signal Processing, and the Computer Societies, the American Association for the Advancement of Science, the SPIE International Society for Optical Engineering, and is in the Who's Who Directory of Optical Scientists and Engineers.

References

1. **Frey KA.** Positron emission tomography. In: Siegel GJ, Agranoff BW, Albers RW, Molinoff PB (Eds): *Basic Neurochemistry: Molecular, Cellular, and Medical Aspects.* New York: Raven Press, Ltd, pp. 839-855, 1989.

2. **Horwitz B, Duara R, Rapoport S:** Intercorrelations of glucose metabolic rates between brain regions: application to healthy males in a state of reduced sensory input. *J. Cereb Blood Flow Metab* 4:484-499, 1984.

3. **Talairach J, Tournoux P:** *A co-planar stereotaxic atlas of a human brain.* Stuttgart, Germany: Thieme-Verlag, 1988.

4. **Friston KJ, Frith CD, Liddle PF, Frackowiak RSJ**: Comparing functional (PET) images: the assessment of significant change. *J Cereb Blood Flow Metab* 11(4):690-699, 1991.

5. **Friston KJ, Frith CD, Liddle PF, Frackowiak RSJ**: Plastic transformation of PET images. *J. Computer Assisted Tomography* 15(4):634-639, 1991.

6. **Brown LG:** A survey of image registration techniques. *ACM Computing Surveys* 24(4):325-376, 1992.

7. **van den Elsen PA, Pol E-JD, Viergever MA:** Medical image matching: a review with classification. *IEEE Engineering in Medicine and Biology* 12(1):26-39, 1993.

8. **Mallat SG:** A theory of multiresolution signal decomposition: the wavelet representation. *IEEE Trans. Pattern Anal. Machine Intell.* PAMI-11(7):674-693, 1989.

9. **Unser M, Aldroubi A.** Polynomial splines and wavelets -a signal processing perspective. In: Chui CK (Ed) : *Wavelets -A tutorial in theory and applications.* San Diego: Academic Press, pp. 91-122, 1992.

10. **Frank J, Verschoor A, Boublik M:** Computer averaging of electron micrographs of 40S ribosomal subunits. *Science* 214:1353-1355, 1981.

11. **Trus BL, Unser M, Pun T, Steven AC:** Digital image processing of electron micrographs: the PIC system II. *Scanning microscopy* Suppl. 6:441-451, 1992.

12. **Unser M, Aldroubi A, Eden M:** The L_2 polynomial spline pyramid. *IEEE Trans. Pattern Anal. Mach. Intell.* 15(4):364-379, 1993.

13. **Aldroubi A, Unser M, Eden M:** Cardinal spline filters: stability and convergence to the ideal sinc interpolator. *Signal Processing* 28(2):127-138, 1992.

14. **Unser M, Aldroubi A, Eden M:** Polynomial spline signal approximations: filter design and asymptotic equivalence with Shannon's sampling theorem. *IEEE Trans. Information Theory* 38(1):95-103, 1992.

15. **Press WH, Flannery BP, Teukolsky SA, Vetterling WT:** *Numerical Recipes.* Cambridge, GB: Cambridge University Press, 1986.

16. **Unser M, Aldroubi A, Eden M:** B-spline signal processing. Part II: efficient design and applications. *IEEE Trans. Signal Processing* 41(2):834-848, 1993.

17. **Unser M, Aldroubi A, Gerfen CR.** A multiresolution image registration procedure using spline pyramids. Proc SPIE Vol. 2034, Mathematical Imaging: Wavelet Applications in Signal and Image Processing: 160-170, 1993.

18. **Ruttimann UE, Unser M, Rio D, Rawlings RR.** Use of the wavelet transform to investigate differences in brain PET images between patients. Proc. SPIE Vol. 2035, Mathematical Methods in Medical Imaging II: 192-203, 1993.

19. **Rioul O, Vetterli M:** Wavelets and signal processing. *IEEE Signal Processing Magazine* 8(4):11-38, 1991.

20. **Daubechies I:** *Ten lectures on wavelets.* Philadelphia, PA: Society for Industrial and Applied Mathematics, 1992.

21. **Rioul O:** A discrete-time multiresolution theory. *IEEE Trans. Signal Proccessing* 41(8):2591-2606, 1993.

22. **Unser M:** On the optimality of ideal filters for pyramid and wavelet signal approximation. *IEEE Trans. Signal Processing* 41(12):3591-3596, 1993.

23. **Battle G:** A block spin construction of ondelettes. Part I: Lemarié functions. *Commun. Math. Phys.* 110:601-615, 1987.

24. **Lemarié P-G:** Ondelettes à localisation exponentielles. *J. Math. pures et appl.* 67(3):227-236, 1988.

25. **Aldroubi A, Unser M:** Families of multiresolution and wavelet spaces with optimal properties. *Numerical Functional Analysis and Optimization* 14(5-6):417-446, 1993.

26. **Unser M, Aldroubi A, Eden M:** A family of polynomial spline wavelet transforms. *Signal Processing* 30(2):141-162, 1993.

27. **Edgington ES:** *Randomization tests.* New York: Marcel Dekker, 1987.

28. **Frank J:** Averaging of low-exposure electron micrographs of non-periodic objects. *Ultramicroscopy* 1:159-162, 1975.