An Interactive Multiscale Framework for Enhancement and Visualization of 2D & 3D Image Data

Andrew D. Graham, Martin J. Turner & Terry Hewitt
Manchester Visualization Centre, Manchester Computing, The University of Manchester, Manchester, UK

Abstract
We present an interactive, multiscale framework for enhancement and de-noising of images and volumes for the purposes of medical image visualization. Among the visualization issues addressed are noise suppression and contrast enhancement. Previous research in this area has concentrated on non-linear operations performed either on the images or their decomposed wavelet coefficients. The vast majority of this research uses global operations over the whole image and sets standard operation parameters for all images. We describe novel non-linear operators and visualization methods that allow the operation parameters to be tailored to an unfamiliar image by the viewer, thereby increasing the adaptability and effectiveness of the enhancement process.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Interaction Techniques

1. Introduction
Analysis and comparison of images produced by medical imaging technologies is an invaluable diagnostic tool for doctors and surgeons. The earliest of these technologies was x-ray radiography, discovered by Wilhelm Conrad Röntgen in 1895. Currently röntgenograms, or radiographs, are still the most commonly used imaging tool but they have been joined by more advanced technologies, such as x-ray computed tomography (CT), magnetic resonance imaging (MRI) and positron emission tomography (PET). There are many challenges in effectively visualizing these types of data. Issues to be addressed include insufficient spatial resolution, noise, clutter and a high dynamic range [MG02].

The issue of resolution in medical images can be demonstrated by examining CT data. A non-helical CT scanner produces a series of sequential, cross-sectional images of the subject. Resolution orthogonal to the image plane is determined by the desired speed of the scan and the maximum x-ray dose for the patient. This generally means that spatial resolution orthogonal to the image plane is lower than that in the image plane. This can be problematic as even the higher resolution in the image plane can be too coarse to visualize some of the tiniest structures of the human body, for example bronchioli in the lungs. As a consequence, the exact shape of small structures cannot be determined. Also, in some cases, a single pixel value can represent the mixed densities of more than one physical structure. This effect is known as partial volume averaging. CT is currently the highest resolution volume imaging technology used in medicine, with MRI second, and PET delivering significantly lower resolution data than either of the other two.

Noise can be defined as artefacts that appear in medical image data that are not representative of physical features of the subject but have been caused by something else. Combined with insufficient resolution, noise can also be problematic. Image features can be as small as one or two pixels in size and therefore are difficult to differentiate from point noise. CT images tend to be relatively low noise due to the scanning method and standard noise reduction algorithms applied to the data inside the scanner itself, but some noise is still present in the images. Other imaging technologies tend to be more affected by noise, with PET producing a particularly low signal-to-noise ratio.

In addition to true noise, physical characteristics the viewer is not interested in can be a barrier to accurate interpretation of image features. This visual clutter is best represented in traditional radiographs, in which there is no depth information and all features throughout the subject appear superimposed, although it also occurs, to an extent, in MRI and CT. Visual clutter can only ever be a subjective assessment, depending entirely on what the viewer is looking for.
Interpretation of medical images is further hampered by the high dynamic range of the data. CT scanners commonly store 12 bits per pixel, allowing for 4096 different gray levels. Most display systems allow 256 gray levels to be displayed. In the worst case, this would force a mapping of 16 similar input values to the same output value, causing low contrast image features to be lost.

This paper presents a framework for interactive enhancement and visualization of medical images. In the following sections we describe and discuss wavelets as a multiscale analysis tool and their suitability for enhancement by non-linear enhancement operations. We then cover the related work in the field of medical image enhancement and de-noising, before describing three new visualization and enhancement techniques and their integration into an interactive framework for enhancement and de-noising tasks by an expert user. We discuss the issues related to enhancement of higher dimensional data and conclude with the presentation and discussion of the results achieved so far.

2. Background

In this section we will cover some of the fundamentals of wavelet transforms, linear and non-linear enhancement and de-noising methods.

2.1. Wavelets

When wavelet transforms are applied to discrete signals, such as images, algorithms are divided into two main types, multi-resolution algorithms [Mal89] and the “à trous” algorithm [MZ92]. Both algorithms transform the data by convolving the input data with digital filters (equations 1,2). The filters are derived, using mathematical methods peculiar to the algorithm, from the definition of a continuous wavelet. Wavelet filters work as pairs, one scaling filter, referred to here as \( h \) and one differencing filter, denoted by \( g \). The output from applying \( h \) to the data is a representation of the original data at a coarser scale. The output from applying \( g \) is the missing detail from the new coarser representation.

\[
Sf(x) = f * h = \sum_{u=0}^{N} f(x)h(x-u) \quad (1)
\]

\[
Wf(x) = f * g = \sum_{u=0}^{N} f(x)g(x-u) \quad (2)
\]

In the case of multi-resolution analysis, the filters maintain a constant size, while the results of the convolutions are down-sampled by a factor of two at each level of decomposition. The result of applying \( h \) and downsampling is to reduce the resolution of the data by half. In standard implementations, the results from the differencing convolution are concatenated onto the results of the scaling convolution. As they have both been downsampled by a factor of two, the result is a representation of the original data that has the same size but is separated into differences and averages. The scaled data can undergo a second pass of the wavelet transform creating an even lower resolution signal with lower resolution differences and so on, hence “multiresolution” analysis.

![Multiresolution Image Transform](image)

This algorithm is applied to 2D images by transforming each scan line with the horizontal 1D filters, then applying the transpose of the filters to each vertical line in the result. This produces four quadrants in the transformed image, one scaled down representation of the original and three wavelet coefficient quadrants. In a multi-level decomposition the upper right quadrant is processed again, as in Figure 1.

Data reconstruction in the multiresolution algorithm is performed by upsampling the scale and difference coefficients and applying reconstruction filters separately, before adding the two results together. Depending on the wavelet satisfying very specific criteria, a perfect reconstruction of the original image can be achieved. Due to the nature of the criteria, there are a relatively small number of known perfect reconstruction wavelets.

In order to allow wavelets to be more adaptable and easier to construct, the “à trous” algorithm sacrifices perfect reconstruction and a constant data size during decomposition. Starting with continuous definitions of the differencing and scaling filters, a reconstruction filter \( k \) can be derived (3), where \( F(\omega) \) is the Fourier transform of \( f(x) \). The filters are applied to the data but, rather than downsampling, all data are kept and the next level of decomposition is performed with filters that have been upsampled by interleaving with zeros. Reconstruction is performed by applying \( k \) to the wavelet data and applying \( \bar{h} \), the complex conjugate of \( h \), to the scaled data, then digitally adding the results.
In the 2D case the difference filters are applied horizontally and vertically to produce two difference images. A single scale image is created by convolving the source image with the matrix constructed by $h \times h^T$.

During the reconstruction phase, the difference images are convolved with the matrix constructed by $k \times k^T$, where $l$ is the filter defined in equation 4. The scale image is convolved with the matrix $h \times h^T$. The three results are added together to produce the reconstructed image.

$$K(\omega) = \frac{1 - |H(\omega)|^2}{G(\omega)}$$

In the 2D case the difference filters are applied horizontally and vertically to produce two difference images. A single scale image is created by convolving the source image with the matrix constructed by $h \times h^T$.

During the reconstruction phase, the difference images are convolved with the matrix constructed by $k \times k^T$, where $l$ is the filter defined in equation 4. The scale image is convolved with the matrix $h \times h^T$. The three results are added together to produce the reconstructed image.

$$L(\omega) = \frac{1 - |H(\omega)|^2}{2}$$

For the application of image enhancement and de-noising, the “à trous” algorithm is considered superior for three reasons. Firstly, and most importantly, the algorithm is translation invariant. This is not true of the multiresolution algorithm because of the downsampling step. Secondly, the algorithm allows near perfect reconstruction with a wide variety of wavelets, enabling the selection of a wavelet that can pick out the features of interest within the image. Finally, because of the simplicity of the filters that can be used, manipulation of wavelet coefficients is a less complicated task.

### 3. Related Work

A large section of work in the field of medical image processing is geared towards solving the above problems by segmenting the structures present in the images in order to enhance/volume render features of interest or to de-accentuate/remove other structures. This review of related work does not cover this area further than to offer a comprehensive survey by Dirk Bartz et al. [BMF03].

In terms of image contrast enhancement, two popular algorithms have been adaptive contrast enhancement [CW98] and adaptive histogram equalization [Sta00]. These kinds of algorithms have limitations that are covered below and in [Jin04]. In order to overcome these limitations, multiscale transforms have been employed. Success has been found with Laplacian pyramids [DSWB02] but more work has focused on wavelet-based enhancement [LJF95, CL99, JFL01, YCB06] which has been proven to contain the Laplacian pyramid as a subset [LJF95].

Laine et al. [LJF95] proposed a formalization to thresholding and non-linear enhancement of wavelet coefficients for enhancement of digital mammography images, which are a low emission form of radiography. They state that linear enhancement is unsuitable for this type of image because highly contrasted features would be enhanced more than low contrast features. Their solution is to use a non-linear thresholding method to suppress small wavelet coefficients for noise removal, while enhancement is performed using a method similar to histogram stretching.

Following on from this work Yinpeng Jin, in his PhD the-
sis [Jin04], detailed a method of applying an adaptive histogram equalization transform to wavelet coefficients at increasing scales. He showed that it is possible to enhance features of a specific scale without accentuating noise that exists at other scales within the image.

These approaches produced good results, the former for digital mammography and the latter for chest CT, but they ignored important shape information that exists both in the original images and the multiscale edge data. In the case of digital mammography the shape information is minimal but there are large structures present in CT images that can be used to remove noise.

In terms of interaction techniques, the Laine paper [LJF95] suggests a particular variety of wavelet that attempts to deal with discontinuous signals and images, which can be used to allow the user to select an area of the image to decompose rather than decomposing the entire dataset.

More recently, an interactive toolbox for wavelet processing is described in [ZM04] that attempts to visualize the wavelet bases alongside the original data and allow basic processing of the wavelet coefficients. The work, performed concurrently to the work in this paper, offers a visualization technique that allows assessment of wavelet response at multiple scales. Enhancement of wavelet coefficients is restricted to thresholding by coefficient strength.

4. Method

This section contains important considerations for wavelet manipulation, followed by details of several visualization and enhancement operations that can be integrated as an interactive framework for wavelet image and volume enhancement.

4.1. Manipulation of Wavelet Coefficients

As stated above, the “à trous” algorithm is preferable for image manipulation. In both of the referenced works [Jin04, LJF95] the wavelets used were derivatives of the Gaussian spline. In brief, the justification for using this group of wavelets is that the scaling filter is a standard image smoothing filter, for example \([\frac{1}{4}, \frac{1}{2}, \frac{1}{4}]\) for the second derivative, and the differencing filters, being derivatives of the Gaussian smoothing function, are, in the case of the first derivative, a good edge detector and in the case of the second, a good feature detector.

Wavelet coefficients produced by this family of wavelets can be positive or negative. It is worth mentioning here two points on coefficient manipulation. In [LJF95], Andrew Laine stated that changing horizontal and vertical wavelet coefficients independently can produce artefacts in the resulting image. For coefficients from the first derivative wavelets he suggests calculating a magnitude and an angle from the horizontal and vertical coefficients, as in (7) and (8).

\[
S = \sqrt{w_{\text{horiz}}^2 + w_{\text{vert}}^2} \tag{7}
\]

\[
P = \arctan \frac{w_{\text{vert}}}{w_{\text{horiz}}} \tag{8}
\]

This allows the coefficients to be manipulated together by some operation \(S' = f(S)\) before reconstruction with \(w_{\text{horiz}} = S' \cos(P)\), and \(w_{\text{vert}} = S' \sin(P)\). As in the case of the multiscale adaptive histogram equalization described in [Jin04] this approach has the effect of treating “up” edges, from dark to light, and “down” edges, from light to dark, in the same way.

In terms of enhancing 2D and 3D image data, an increase in the magnitude of a wavelet coefficient will be reflected in the resultant image, produced by the wavelet reconstruction process. If coefficients related to noise in the image are reduced then the reconstructed image will appear less noisy. Equally, if coefficients representing slightly contrasted image features are enhanced, the image features themselves will be enhanced.

4.2. Morphological Wavelet Processing

Applying grayscale morphological operators directly to an image, in order to suppress or enhance features, has a destructive effect on image features at a smaller scale than the structuring element. While this is not a problem when applying morphology to the smallest features in an image using small structuring elements, such as in the case of point noise suppression, if we are to apply morphology to larger scale image features we must include it in a framework that protects the smaller image features. Employing morphological operations within the multiscale wavelet framework is a potential solution to this problem.

As in our other wavelet operations, we iteratively decompose the image into its wavelet coefficients at increasing scales. Morphological operations can be applied to the wavelet coefficients at any scale before the image is reconstructed to produce the enhanced result. The interactive nature of the structuring element, morphological operation, location and scale selection allow a great deal of freedom in terms of which image features are enhanced or suppressed.

In Figure 2, we can see the result of applying grayscale closing with a ball shaped structuring element to the second, third and fourth level wavelet coefficients. This has the effect of accentuating the (white) blood vessels in the lung and darkening the other areas. Notice that the technique does not alter the position of the blood vessel boundaries.
4.3. Coefficient Contribution Visualization (CCV)

Wavelet based enhancement and de-noising methods on unfamiliar images suffer the disadvantage that the scale response of the image is unknown. While it is possible to say that a wavelet decomposition will produce strong wavelet coefficients for image details at a particular scale, it is not possible to anticipate, from viewing of the image alone, which areas of the image will contain details of which scale. Without this information, it is unclear at which scale noise or features of interest occur and hence, which levels of the decomposition to enhance or suppress.

The contribution of wavelet coefficients at a particular scale to image features or noise can be visualized by decomposing the image, increasing the magnitude of the coefficients of that scale by a uniform amount and reconstructing the image. When the original image is digitally subtracted from the reconstructed image, the remainder represents the contribution of the coefficients at that scale to the whole image. In order to compare and contrast the contribution of one scale against another, the contributions from three separate scales are assigned to the primary colours of a single, colour result.

In Figure 5, we show the Lenna image and its CCV. The image was decomposed using the second derivative Gaussian spline wavelet defined in [LJF95]. The colour mapping was chosen based on the frequency of light. The large scale coefficients, analogous to low frequency in a Fourier transform, are represented with red, smaller scale contributions are green and the smallest scale is blue.

It is easy to see from the CCV that different image features respond strongly to different scales of the wavelet transform, thus enabling them to be suppressed or enhanced to the exclusion of features at another scale. Where there is a strong response at all scales the pixel colour appears as a combination of all components. Such is the case with the brim of Lenna’s hat, which has a crisp edge but is also a major feature, detectable at larger scales. This results in the area showing gray/white.

In the context of an interactive framework for visualization of images, particularly medical images, the CCV can allow the viewer to tailor the enhancement and de-noising operations below in terms of spatial and scale localization in order to achieve a better enhancement of any image than could be achieved by performing the operations using global settings.

4.4. 2.5D Multiscale Adaptive Histogram Equalization

The algorithm described in [Jin04] for multiscale adaptive histogram equalization (MAHE) of CT images enhances low contrast features in the image by wavelet deconstruction of the image, adaptive histogram equalization (AHE) on the wavelet coefficients and reconstruction. The non-linear AHE operation is redefined as

\[
AHE(I_{x,y}, s) = \frac{\text{count}(I_{x-sx+iy-tx-ty} < I_{x,y})}{\text{count}(I_{x-sx+iy+tx+ty})}
\]

where \(I_{x,y}\) is an image pixel and \(s\) is an offset used to define an area surrounding the pixel. In words, each pixel is set to the count of all pixels in the area with a lower value than the centre pixel divided by the total number of pixels in the area. When applied to a CT image, the AHE algorithm tends to accentuate noise in the image. MAHE avoids this by not enhancing the smallest scale coefficients from the first level of the wavelet decomposition, because image noise tends to be small scale.

The limitations of the MAHE algorithm for CT image enhancement are that it tends to accentuate large image edges by much more than small edges. This is because these edges, like the brim of Lenna’s hat, are evident in every level of the wavelet decomposition and, during enhancement, are also enhanced at every level. This becomes problematic because small scale features that are also low contrast can easily be lost during visualization because of the high dynamic range of the result.

In order to prevent this, our technique takes into consideration the neighbourhood of CT image slices around the slice being enhanced. Since CT scan data are often consecutive axial slices of a subject with a small interceding space between each slice, we can assume that many of the same features will exist in the local slice neighbourhood that exist in the image we are interested in enhancing. In the case of a strong edge in the image, which the 2D MAHE algorithm would tend to over-enhance, the slices above and below current slice will tend to contain the same edge. If we consider those pixels, which are as large or larger than the edge we’re enhancing, as part of the adaptive histogram equalization then the numerator in equation (9) will be reduced and the edge less strongly enhanced. The opposite will be true of fine detail in the image, which is less likely to be surrounded by strong edges in surrounding slices and hence should be more strongly enhanced.

We perform the 2D wavelet decomposition of all slices in
the local neighbourhood and reformulating the AHE step as a 3D operation (10).

\[
AHE(V_{x,y,z}, s, d) = \frac{\text{count}(V_{x-kx+kx+ky+kz-lz+t} < V_{x,y,z}, l < V_{x,y,z})}{\text{count}(V_{x-kx+kx+ky+kz-lz+t})}
\]

(10)

In (10), the z-axis offset \( l \) can be defined by \( l = s/d \). The xy offset, \( k \), changes depending on the z offset and can be defined as \( k = |z - z_{\text{current}}|/d \), for \( z - l \leq z_{\text{current}} \leq z + l \). This reduces the in-plane area as the offset from the CT slice of interest increases, describing a 3D diamond shape in the data. As an example, for values \( s = 5, d = 1 \), the algorithm would include the \( 11 \times 11 \) square around \( V_{x,y,z} \) in the image of interest, the \( 9 \times 9 \) squares in the images above and below, the \( 7 \times 7 \) squares in their immediate neighbours and so on.

The diamond shape of the 3D AHE algorithm is more suitable to the problem than a cube shape as we wish the grayscale information from the image of interest to factor strongly in the result, with slices further away from the image of interest making smaller contributions. The introduction of the variable decay parameter \( d \) introduces the potential to increase the rate of decay of influence of the surrounding slices, flattening the diamond. This would be beneficial in areas where there is a lot of variation orthogonal to the image plane or in a case where the CT data had been acquired with larger spaces between slices. As \( d \) tends towards infinity, 2.5D MAHE behaves similarly to the original MAHE algorithm.

A “true 3D” MAHE algorithm could be implemented by an extension of the wavelet transform from 2D to 3D. This would not be appropriate with non-helical CT scanners as the data produced is not itself truly 3D, rather a concatenation of discrete images sampled separately, but could potentially be used with MRI or helical CT data.

5. Initial Results and Discussion

This section covers some of the initial experimental results for the above techniques.

5.1. Visualization Results

The work in [ZM04] follows a similar line to our own CCV method in that three scales are represented in the same image by assigning primary colours to scales. The acquisition of the colour values is simpler as the raw wavelet coefficients from each scale decomposition are used. Using this method with any other wavelet than the Haar wavelet from their example generates an inaccurate impression of the likely results from manipulating the wavelet coefficients because it ignores the reconstruction filter.

In Figure 6, the left image is the wavelet coefficients of a chest CT slice visualized according to the algorithm in [ZM04]. On the right is the same slice visualized using the CCV method described above. Both images use blue for second level, green for third level and red for fourth level wavelet decomposition coefficients. It can be seen that the left image gives shows an obvious partition in scales between features of the image. This visualization gives the viewer the impression that changing the coefficients at one of these scales will only affect the areas of the image indicated by the areas of the visualization with that colour. It can be seen from the CCV that this is not the case. Rather than the areas in the lungs (the two large roughly symmetrical circular structures) being largely represented by second and third level coefficients they are, in fact, equally present in all levels of the decomposition as indicated by their grey/white colouring in the CCV. Further to this, enhancement of the fourth level coefficients (red) will effect a much wider area around the major image features than suggested by the Zhang method, causing a halo effect in any enhancement.

While the Zhang method may appear cleaner than the CCV, the CCV gives a truer picture of the likely outcomes of coefficient manipulation.

5.2. MAHE Comparison

Figure 3 shows a comparison of a section of a single scan line from an image (a) that has been enhanced first with 2D MAHE (b) and then with our 2.5D MAHE (c). The MAHE enhanced scan line has large features that have been enhanced by a large amount, while some smaller features have been enhanced by less. A comparison with our 2.5D MAHE result shows that the large features have been enhanced by less without significantly reducing the enhancement of the small features.

In Figure 4 we can see the effect this has on the resulting image. The MAHE image has very bright, large features
that span 3000 gray levels. As a result of this high dynamic range, smaller features with less enhancement, such as the smaller features in the lungs, are less observable. In the 2.5D MAHE image, the large structures have been enhanced by far less and the number of gray levels has been reduced to 2000 without a substantial decrease in the enhancement of the small structures.

6. Conclusions and Further Work

We have shown a method of integrating the “à trous” wavelet transform, a novel visualization method and novel local or global non-linear enhancement techniques into an interactive image enhancement and visualization package. We have shown promising initial results for the non-linear techniques and the idea of interactive wavelet visualization as a whole.

Further work will concentrate on extending the current suite of non-linear enhancements, in particular the extension into full 3D for all techniques. Some work must be done to find an effective way to visualize the volume of wavelet contributions for a 3D dataset effectively.

References


Figure 4: Left: MAHE, from [Jin04] of a Chest CT image; Right: 2.5D MAHE of the same image.

Figure 5: Lena Image and its Coefficient Contribution Visualization

Figure 6: Left: Wavelet coefficient visualization from [ZM04] of a Chest CT image; Right: CCV of the same CT image.