Chapter from the book
Downloaded from: http://www.intechopen.com/books/

Interested in publishing with IntechOpen?
Contact us at book.department@intechopen.com
Wavelet Analysis for the Extraction of Morphological Features for Orthopaedic Bearing Surfaces

X. Jiang, W. Zeng and Paul J. Scott
University of Huddersfield, United Kingdom

1. Introduction

Surface texture is one of the most critical factors and important functionality indicators in the performance of high precision and nanoscale devices and components. The functions that have been identified in various studies include wear, friction, lubrication, corrosion, fatigue, coating, paintability, etc. [1-3]. It is also reported that the wear rates of surfaces in operational service is determined by roughness, waviness and the multi-scalar topographic features of a surface, such as random peaks/pits and ridges/valleys. These functional topographical features will impact directly on wear mechanics and physical properties of a whole system, such as hip joint replacement system in bioengineering [4-9]. For example, during functional operation of interacting surfaces, peaks and ridges will act as sites of high contact stresses and abrasion; consequently wear particles and debris will be generated by such surface topographical features, whereas the pits and valleys will affect the lubrication and fluid retention properties. In this situation, a vitally important consideration for functional characterisation must be the appropriate separation of the different components of surfaces, which is not only to extract roughness, waviness and form error, but should also be extended to all multi-scalar topographical events over surfaces.

Effectively surface characterization can help to: (1) Judge whether the manufacturing process or manufacturing conditions are effective or out of control, i.e. whether specific events in the manufacturing processes have occurred; (2) Interpret functional properties of macro, micro and nano geometry, which reflect product properties, such as optical quality, tribological properties, service life, safety, reliability, etc.

Random process techniques have been applied to surface analysis in which the surface topography is assumed as a stationary random system. However, the disadvantage of random analysis is that significant events involved in a surface, such as large peaks/pits, or ridges/valleys will be smoothed over in the Fourier space of a signal, without indicating the location of the frequency events. The early studies of morphology features of surface topography only considered the roughness component of the surface texture. Comprehensive studies of surface characterisation have nearly always concentrated on roughness assessment and have mainly developed evaluation techniques based on parameters and filtration techniques. This is normally accomplished by using Fourier Transform based filtering techniques, especially Gaussian filtering. These techniques are employed to isolate the roughness/waviness frequency bands relevant to the surface, by
breaking down a surface signal into a series of harmonics waves. In fact, Gaussian filtering, used to extract the roughness and waviness components, is based on two presuppositions. One presupposition is that before filtering, the irrelevant form and translation errors have been removed from the measured data set. The other is that the residual surface, obtained according to the first assumption, can be broken down into a series of harmonic components. If surfaces conform to these assumptions, the roughness and waviness can be identified well and extracted from the original surfaces with a modified amplitude resulting from the transmission characteristics of a Gaussian function. These types of filtering techniques only considered the frequency information and do not consider the spatial information of the features.

The Fourier transform has a very good localization property in the frequency domain, however, it lacks the ability to localize in the time/space domain. Figure 1 gives a typical example of the Fourier transform of two different signals with the same sine wave components, while one is stationary and the other is non-stationary. The two signals have the same Fourier transform spectrum. To distinguish between these two signals, the analysis tools need the time/space-frequency localization property. There are two functions that link the time/space-frequency domain, the ambiguity function (in radar applications) and the Wigner distribution (in quantum mechanics). Both have been introduced to identify the topographic features of engineering surfaces by Whitehouse and Zheng [10, 11]. Among these tools, the Wigner distribution are more suitable as it preserves the absolute space and frequency parameters [10]. The energy distribution in the space-frequency plane, offered by the Wigner transform, can be used to identify the variation in surface topography. The result is that the information about the frequencies and their locations in a surface signal has been used to monitor and adjust the manufacturing processing. However, it is also well known that although the Wigner transform can offer an analysis of the energy distribution of a signal, it only yields an imperfect description of the energy distribution, due to its substantial interference terms.

During the last decade, novel filtering techniques based on robust Gaussian regression and multi-scale techniques have been developed. Robust filtering has been designed to eliminate problems in the evaluation of the roughness of multi-process surfaces such as plateau honed cylinder bores and is critical to the implementation of areal surface analysis techniques. As previously stated, analysis of multi-scalar properties of a surface topography needs to not only provide both the frequencies of the signal and their location, but also to accurately recover and perfectly reconstruct these topographic features. Considering these needs of surface assessment, wavelet analysis has been applied to surface characterization [12-16]. Wavelet analysis employs time-frequency windows and offers the relevant time/space-frequency analysis. The Wavelet transform has been proved to be a powerful tool for various applications, for example, the Wavelet series expansions, developed for applied mathematics and signal processing. Multi-scalar features have been used in capturing, identifying and analyzing local, non-stationary processes; Sub-band coding used in encoding, compressing, reconstructing and modelling signals and images [17-20].

In this chapter three generations of biorthogonal wavelet transforms, for the extraction of morphological structures from micro/nano scalar surfaces in the field of bioengineering, will be introduced. The chapter’s aim is to create a “tool box” of wavelet techniques capable of complex analysis and interpretation of surface topography data; leading to the extraction of functionally critical morphological features from micro/nano scalar surfaces of orthopaedic joints for in-vitro and clinically retrieved applications.
Wavelet Analysis for the Extraction of Morphological Features for Orthopaedic Bearing Surfaces

(a) a stationary signal composed of sine waves $\{30\text{Hz}, 45\text{Hz}, 60\text{Hz}\}$

(b) a nonstationary signal composed of sinewaves $\{30\text{Hz}, 45\text{Hz}, 60\text{Hz}\}$

(c) Same Fourier spectrum of (a) and (b)

Fig. 1. Fourier transform of two different signals with the same sin components
2. Wavelet

A Wavelet \( \psi(t) \) is a waveform that has compact support in both the space and frequency domains and whose integral is zero. Compare wavelets with sine waves, which are the basis of Fourier analysis. Sinusoids do not have limited duration — they extend from minus to plus infinity, where sinusoids are smooth and predictable, early wavelets tended to be irregular and asymmetric. The wavelet \( \psi(t) \) function must satisfy the admissibility condition:

\[
C_v = \int_{-\infty}^{\infty} \left| \hat{\psi}(\omega) \right|^2 d\omega < \infty
\]

which guarantees the existence of the inverse wavelet transform.

In wavelet analysis the signal is broken down into rescaled and shifted versions of the original waveform. This transfers space based information into scale based information, which represents the frequency and location properties of the original signal. The wavelet transform can be defined as [17-18]:

\[
W_{a,b}(f) = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) dt
\]

where, \( f(t) \) is the signal, \( a \) and \( b \) are the dilation and translation parameters for the function \( \psi(t) \), which is a wavelet function or mother function. The wavelets, \( \psi_{a,b}(t) \), (basis functions) can be obtained by dilation and translation of the mother wavelet \( \psi(t) \). The parameter \( a \) can be used to illustrate dynamic transmission bandwidths or cover different frequency ranges, and \( b \) is the translation parameter that can be used to define the location of signal events. Typically one-dimensional continuous wavelets are expressed by [17-18]:

\[
\psi_{a,b}(x) = \frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a}\right)
\]

In the real world, most signals are discrete and the dilation and translation parameters both taking only discrete values. The rescaled wavelets, \( \psi_{a^j,b}(t) = \psi(a^{-j}t) \) are dilated by the factor \( a^{-j} \), this is illustrated in figure 2(a) where the parameter \( a \) takes the discrete value of two. \( j \) is the scalar parameter that can be used to illustrate dynamic transmission bandwidths or cover different frequency ranges. The shifted wavelets \( \psi_{a,k}(t) = \psi(t-k) \) are translated by \( k \) (translation parameter) as shown in figure 2(b). By changing \( k \), the time location centre has been moved. Typical one-dimensional wavelets \( \psi_{a,k}(x) \) are dilated \( j \) times and shifted \( k \) times. The discrete wavelets are expressed by:

\[
\psi_{j,k}(x) = a^{-j/2} \psi(a^{-j}x - k)
\]

Thus a signal can be divided into different scales and different locations, with the signal data now being represented on a 'time-scalar plane'. Multi-resolution divides the frequencies into octave bands, from \( \omega \) to \( 2\omega \). Frequencies shift upward by an octave when
Wavelet Analysis for the Extraction of Morphological Features for Orthopaedic Bearing Surfaces

Fig. 2. Example of a discrete wavelet

time is rescaled by two. Figure 3(c) shows how the time-frequency plane is partitioned naturally into rectangles of constant area.

For a discrete signal \( z(x) \in L^1(\mathbb{Z}) \), its discrete wavelet transform can be expressed by:

\[
W_{j,k}(f) = \langle f(x), \psi_{j,k}(x) \rangle = 2^{-j/2} \sum_{k} f(x) \psi(2^{-j} x - k)
\]  

(5)

The signal \( z(x) \) can be recovered with the following equation:

\[
z(x) = \langle z(x), \psi_{j,k}(x) \rangle \psi_{j,k}(x)
\]  

(6)

where, the \( \Psi_{j,k} = (F^*F)^{-1} \psi_{j,k} \).

A basic idea behind wavelet analysis is the transfer of a complex frequency analysis into a simple scalar analysis. Wavelet analysis employs time-frequency windows and offers the relevant time-frequency analysis, which uses long windows at low frequencies and short windows at high frequencies (as shown in figure 3c). Wavelet mathematics provides a mathematical microscope that “can divide functions into different frequency components, and then study each component with a resolution that is matched to its scale” [17-18]. The wavelet transform is an update of the classical Short-Time Fourier Transform (STFT) or Gabor Transform as well as the Wigner distribution [21-22]. In contrast, STFT uses a single size analysis window (shown in figure 3b), and offers the same accuracy of analysis in the
whole time-frequency plane. The Wavelet transform is related to space-frequency analysis and is similar to the Wigner-Ville distribution, but is better able than the STFT to "zoom in" on very short-lived high frequency phenomena, such as transients in signals or singularities in functions. A point to be noted is that wavelet analysis takes the signal decomposition to a space-scalar (time-frequency) plane and separates then reconstructs these components in the space domain. In contrast, Wigner analysis also decomposes a signal into the space-frequency plane, then studies the energy distribution of these components but cannot reconstruct the modified signal perfectly.

Fig. 3. The different windows of the FFT, STFT and WT

The Wavelet technique includes many different wavelet functions: each has its own properties and applications. Linear phase (symmetric wavelet) and finite pulse filtering are basic properties in surface analysis, only the Haar wavelet and the biorthogonal wavelet have both characteristics. Due to the fact that the Haar wavelet is binary and a discontinuous function makes it not suitable for surface analysis, only the biorthogonal wavelet are useful tools for surface decomposition.

3. The Bi-orthogonal wavelet for surface analysis

The main approach to the discrete wavelet transform is to use a two-channel filter bank. The dilation equations of the basis wavelets and the corresponding scaling functions have associated lowpass $H_0$ and highpass $H_1$ filter coefficients.

The basic property of the discrete wavelet function and the scale function for multi-scale analysis are their two-scale difference function [17-19]:

$$\begin{align*}
\psi(x) &= \sum_{k} 2h_0(k)\phi(2x - k) \\
\phi(x) &= \sum_{k} 2h_1(k)\phi(2x - k)
\end{align*}$$

where, $\psi(.)$ is the wavelet and $\phi(.)$ is the scale function. The two-scale difference functional relationships means that the wavelet function and the scale function at scale $j$ can be represented as a linear combination of the scale functions at scale $j+1$. Here, the weighting coefficients $h_0(k), h_1(k)$ are the impulse response of the lowpass filter $H_0$ and highpass $H_1$ respectively. For orthogonal wavelet transforms $H_0$ and $H_1$ are orthogonal. However, the frequency responses of the two-channel filters, the so-called analysing filters, are not ideal brick wall filters since normally there are overlaps that would lead to aliasing, amplitude modifications and phase distortions.
3.1 The first generation biorthogonal wavelets

The biorthogonal wavelet overcomes some of the defects of the orthogonal wavelet and other two-channel filters. In a biorthogonal wavelet the forward \((H_0 \text{ and } H_1)\) and inverse filters \((G_0 \text{ and } G_1)\) are different. The inverse filters \(G_0\) and \(G_1\) (the so-called synthesis filters), are specially designed to compensate for these errors of the analysing filters \(H_0\) and \(H_1\). When the frequency responses of the synthesis filters, \(G_0\), \(G_1\), are the inverses of the analysis filters, \(H_0\), \(H_1\), they satisfy [23]:

\[
G_0(z)H_0(z) + G_1(z)H_1(z) = 2z^{-l} \\
G_0(z)H_0(-z) + G_1(z)H_1(-z) = 0
\]

(8)

the errors in this analysis bank are cancelled. Where, \(l\) is time delay, the distortion term must be \(z^{-l}\), and the aliasing term must be zero.

In the biorthogonal case, abandoning the orthogonal relationship between the lowpass analysis filters \(H_0(z)\) and the highpass analysis filter \(H_1(z)\), while keeping the biorthogonal conditions between \(H_0(z), G_0(z)\), and \(H_1(z), G_1(z)\). The biorthogonal condition can be expressed by the impulse responses, \(h_0(k), h_1(k), g_0(k), g_1(k)\), of \(H_0(z), H_1(z), G_0(z), G_1(z)\), as:

\[
\sum h_0(k)g_1(k-2n) = 0 \\
\sum h_1(k)g_0(k-2n) = 0
\]

(9)

Where, \(\tilde{g}_0\) and \(\tilde{g}_1\) are the time reverse of \(g_0\) and \(g_1\) respectively. Formulae (9) means \(h_0(k)\) is orthogonal to \(\tilde{g}_1(k-2n)\) and \(h_1(k)\) is orthogonal to \(\tilde{g}_0(k-2n)\). Here, \(\tilde{g}_0(k-2n)\) and \(\tilde{g}_1(k-2n)\) are the even shift \(2n\) of \(\tilde{g}_0(k)\) and \(\tilde{g}_1(k)\) respectively.

According to Wavelet theory [17-18], the dilation equations of the analysis wavelet equation, \(\psi(x)\), and synthesis wavelet function, \(\psi(x)\) can be constructed from the impulse responses of the analysis and synthesis highpass filters respectively [23]:

\[
\begin{aligned}
\tilde{\psi}(x) &= \sum_{l=0}^{N} 2h_1(l)\tilde{\varphi}(2x-l) \\
\psi(x) &= \sum_{k=0}^{N} 2g_1(k)\varphi(2x-k)
\end{aligned}
\]

(10)

the scale equation are given by:

\[
\begin{aligned}
\tilde{\phi}(x) &= \sum_{l=0}^{N} 2h_0(l)\tilde{\varphi}(2x-l) \\
\phi(x) &= \sum_{k=0}^{N} 2g_0(k)\varphi(2x-k)
\end{aligned}
\]

(11)

The \(\tilde{\phi}(x), \tilde{\psi}(x), \varphi(x), \psi(x)\) are constructed as a biorthogonal wavelet pair, their limit functions would inherit biorthogonality.

For areal surface, the analysis and synthesis scale functions, \(\tilde{\phi}(x,y)\) and \(\phi(x,y)\) can be expressed by:

\[
\tilde{\phi}(x,y) = \sum_{l=0}^{N} 2h_1(l)\tilde{\varphi}(2x-l)
\]

\[
\phi(x,y) = \sum_{k=0}^{N} 2g_1(k)\varphi(2x-k)
\]

(12)
\[ \phi(x, y) = \phi(x) \phi(y) \]
\[ \phi(x, y) = \phi(x) \phi(y) \]

Using the scale function pair, the analysis wavelets and synthesis wavelets are built by [27]:

\[
\psi(x, y) = \begin{cases} 
\psi^h(x, y) = \phi(x) \psi(y) \\
\psi^v(x, y) = \psi(x) \phi(y) \\
\psi^d(x, y) = \psi(x) \psi(y) 
\end{cases}
\]
\[ \psi(x, y) = \begin{cases} 
\psi^h(x, y) = \phi(x) \psi(y) \\
\psi^v(x, y) = \psi(x) \phi(y) \\
\psi^d(x, y) = \psi(x) \psi(y) 
\end{cases} \tag{13} \]

where, \( \psi^h(x, y) \), \( \psi^v(x, y) \) and \( \psi^d(x, y) \) give the analysis wavelets in the horizontal, vertical and diagonal directions, respectively. \( \psi^h(x, y) \), \( \psi^v(x, y) \) and \( \psi^d(x, y) \) give the synthesis wavelets of the other three directions.

By employing a 2D biorthogonal wavelet pair, an \( N \times N \) points areal discrete surface signal \( z(x, y) \) can be transformed to a wavelet series by decomposing the signal data \( z(x, y) \) on the 2D biorthogonal wavelets basis:

\[ W_{j,k}(z) = \langle z(x, y), \psi_{j,k}(x, y) \rangle = \langle d_{j,k1}, ..., d_{j,kJ}, a_{j,k} \rangle \tag{14} \]

The \( a_{j,k} \) are approximation coefficients of the signal \( z(x, y) \) at the scale \( 2^{-j} \), and the \( d_{j,k1}, ..., d_{j,kJ} \) are a group of detail coefficients of \( z(x, y) \) at the scalar range of \( 2^{-j} \) \( (j = 1 \sim J, J \) is the maximal wavelet decomposition Level., and \( k \) refers to a subset \( (1 : \frac{N}{2^j}) \) ). The \( a_{j,k} \) are generated by the set of inner products [27]:

\[ a_{j,k} = \langle A_{j-1}(x, y), \phi_{j,k}(x, y) \rangle \tag{15} \]

\( A_{j-1}(x, y) \) is called a discrete scalar approximation of \( z(x, y) \) at the scale \( 2^{-j} \), and represents the low frequency components of \( z(x, y) \) at the scale \( 2^{-j-1} \). When the scale \( j \) is set to zero, the scalar approximation \( A_0(x, y) \) equals \( z(x, y) \). \( A_j(x, y) \) at the scale \( 2^{-j} \) can be reconstructed by using the set \( a_{j,k} \) \( (j = 1 \sim J, k \in (1 : \frac{N}{2^j}) \) ), and applying the inverse wavelet transform:

\[ A_j(x, y) = \sum_k a_{j,k} \phi_{j,k}(x, y) = \sum_k \langle A_{j-1}(x, y), \phi_{j,k}(x, y) \rangle \phi_{j,k}(x, y) \tag{16} \]

The detail coefficient, \( d_{j,k} \), is a combination of three directional detail parts:

\[ d_{j,k} = \begin{cases} 
d^{h}_{j,k} = \langle A_{j-1}(x, y), \psi_{j,k}^h(x, y) \rangle \\
d^{v}_{j,k} = \langle A_{j-1}(x, y), \psi_{j,k}^v(x, y) \rangle \\
d^{d}_{j,k} = \langle A_{j-1}(x, y), \psi_{j,k}^d(x, y) \rangle 
\end{cases} \tag{17} \]

Similarly, the inverse wavelet transform of the detail coefficient \( d_{j,k} \) \( (j = 1 \sim J, k \in (1 : \frac{N}{2^j}) \) gives the high frequency components of \( z(x, y) \) at the scale \( 2^{-j} \) and it is called a discrete...
detail $D_j(x,y)$ of $z(x,y)$ at the scale $2^j$:

$$D_j(x,y) = D^a_{j,k}(x,y) + D^v_{j,k}(x,y) + D^d_{j,k}(x,y)$$

$$= \sum_k d^{a}_{j,k}\psi^a(x,y) + \sum_k d^{v}_{j,k}\psi^v(x,y) + \sum_k d^{d}_{j,k}\psi^d(x,y)$$  \hspace{1cm} (18)

The equations show that in the two-dimensional case, the approximation and detail coefficients are computed by separate filtering of the signal $z(x,y)$ along the abscissa and ordinate. The wavelet decomposition can thus be interpreted as a signal decomposition in a set of independent, spatially oriented frequency channels. The approximation coefficient $a_{j,k}$ is decomposed into $a^a_{j,k}$, $a^v_{j,k}$, $d^v_{j,k}$, and $d^d_{j,k}$. The coefficients $a_{j,k}$ corresponds to the lower frequencies, the coefficients $d^a_{j,k}$ correspond to the vertical high frequencies (horizontal edges), the coefficients $d^v_{j,k}$ the horizontal high frequencies (vertical edges) and the coefficients $d^d_{j,k}$ the high frequencies in both directions. Consequently, the areal surface $z(x,y)$ can be reconstructed using:

$$z(x,y) = A_1(x,y) + D_1(x,y) = A_1(x,y) + D_2(x,y) + D_1(x,y)$$

$$= \ldots = A_1(x,y) + D_j(x,y) + D_{j-1}(x,y) + \ldots + D_1(x,y)$$ \hspace{1cm} (19)

### 3.2 The second generation biorthogonal wavelet

The fundamental idea behind the use of Biorthogonal wavelet filtering for surface analysis is to break down areal surface raw data set into a rescaled and shifted version of the original waveforms in a ‘scalar space’, extract the waveforms carrying different information and then reconstruct directly these components. The main advantage of using a biorthogonal wavelet is that it has a linear phase (leading to real output without aliasing and phase distortion) and a traceably located property so that the different component surfaces obtained can more naturally record the real surface. However due to the fact that wavelets (basic functions) are built by dilation and translation of the prototype wavelet which relied on the Fourier transform, and the fact that the Wavelet transform needs to be applied along three directions, horizontal, vertical, and diagonal, the theory and corresponding algorithm are very complex. Furthermore, there is still the boundary destruction inherent when using the Fourier transform.

For engineering application purposes the method used for surface characterisation must be simple and natural. In other words, the method of the separation and reconstruction of different components of the surface should have both the ‘simplicity’ of general surface filter and ‘naturalness’ of a biorthogonal wavelet filter, as result of this, the second generation of biorthogonal wavelet, originally developed at Bell Laboratory [24-26], was introduced into surface analysis. The wavelet and scalar coefficients are only dependant on the measured raw data of a surface and the filtering and lifting factors calculated by a cubic spline interpolation in an interval (Dyn et al. 1987; Flowers 1995; and Stoer 1980). Compared to the first generation of Wavelet representation, the new model does not take the Fourier transform as a prerequisite. The Wavelet transform only embraces three stages, splitting, prediction and updating. The other advantage derived from this model is that there is no
boundary destruction. Although the implementation of the lifting wavelet is completely different from the former model, it is much easier to understand and perform.

In the lifting scheme, the splitting procedure is to divide an original signal into even and odd subsets. The even subset is obtained by subsampling the input signal at the even positions, and it can be seen as the initial approximation of the input signal. The odd subset is obtained by subsampling the input signal at the odd positions. The lost information of the initial approximation is contained in the odd subsets. In the prediction step: the even subset is used to approximate the odd subset at the corresponding position by using interpolation on the neighboring data. Then the difference between the odd subset and the approximated subset can be regarded as the output of the high pass filter, or the wavelet coefficients at this specified level. Finally, in the updating step: the final approximation of the input signal can be calculated by updating the even subset (initial approximation of the input signal) with the help of neighboring wavelet coefficients obtained from last step. The output of this step is the approximation coefficients of low-pass filter. The benefit of using the lifting scheme is: (1) the lifting scheme allows a fully in-place calculation of the wavelet transform; (2) it is very simple as the inverse wavelet transform with the lifting scheme is only a reversal of the operations of the forward transform.

In this implementation, the analysis highpass filter, $H_1$, and synthesis lowpass filter, $G_0$, of the initial finite biorthogonal wavelet filter set, $\{H_0, H_1, G_0, G_1\}$ within the first generation are transferred to $H_1, G_0$ which can be found by the lifting scheme as\([24-26]\):

$$
\begin{align*}
H_1(z) &= H_0(z) + G_0(z)S(z^2) \\
G_0(z) &= G_0(z) - H_0(z)S(z^2)
\end{align*}
$$

(20)

where the $S(z)$ is a Laurent polynomial. Substituting this new set to the equation (7), in order to perfectly reconstruct the signal, the frequency responses of analysis and synthesis filters of the second algorithm should also satisfy\([24-26]\):

$$
\begin{align*}
G_0^2(z)S(z^2) - H_0^2(z)S(z^2) &= 0 \\
G_1(z)G_0(-z)S(z^2) - H_0(z)H_0(-z)S(z^2) &= 0
\end{align*}
$$

(21)

In this case, $S(z^2) = S(z^2)$. After lifting has been performed, the new biorthogonal wavelet pair can be found as:

$$
\begin{align*}
\tilde{\psi}'(x) &= \tilde{\psi}(x) - \sum_{l=0}^N x(l)\tilde{\phi}(x-l) \\
\psi'(x) &= 2\sum_{k=0}^N g_1(k)\phi'(2x-k)
\end{align*}
$$

(22)

and

$$
\begin{align*}
\tilde{\phi}(x) &= 2\sum_{l=0}^N h_0(l)\tilde{\phi}(2x-l) \\
\phi'(x) &= 2\sum_{k=0}^N g_0(k)\phi'(2x-k) + \sum_{k=0}^N s(k)\psi'(x-k)
\end{align*}
$$

(23)
where the coefficients, \(s(\cdot)\), are from the Laurent polynomial \(S(z)\), the order is based on how many neighbouring coefficients are chosen to interpolate in the prediction and updating steps. The power behind the lifting scheme is that \(s(\cdot)\) can be used to fully control all wavelets and synthesis scaling functions.

The simplest wavelet transform using lifting scheme is when \(N = \bar{N} = 0\), which means in the predict and updating step, no neighbouring coefficients are used to interpolate and update. In this case, the wavelet transform does nothing but separates the input data into even and odd subset, which is called lazy wavelet.

In the lifting wavelet transform with \(N = 2\) the coefficients, \(s(l) = \left(\frac{1}{2}, \frac{1}{2}\right)\), the wavelet coefficient \(d_{j,k}\) encodes the difference between the exact sample \(A(0,2j+1)\) and its linear prediction of two even neighbours \(A_{0,2k}, A_{0,2k+2}\). The wavelet coefficients can be written as:

\[
d_{j,k} = A_{j,2k+1} - 1/2(A_{j-1,2k} + A_{j-1,2k+1})
\]

Employing the dual lifting scheme, \(\bar{N} = 2\), that is, the coefficient \(s'(k) = \left(0, 0, 1, 1\right)\), the scalar coefficients \(a_{j,k}\) of \(A_{0,2k}\) would be updated by wavelet coefficients \(d_{j,k}\) and \(d_{j,k+1}\). The scalar coefficients can be expressed as [24-26]:

\[
a_{j,k} = A_{j,2k} + 1/4(d_{j,k} + d_{j,k+1})
\]

### 3.3 The lifting wavelet algorithm

Cubic spline interpolation has been used and verified as an appropriate Laurent polynomial, \(S(z)\) in the prediction stage since the resulting wavelet satisfies the requirements of surface analysis: excellent refinement accuracy, perfect reconstruction, minimum sampling condition (measured area) and a minimum computation.

The three stages of the Wavelet transform by the lifting scheme are depicted in figure 4. The fast lifting wavelet transform algorithm is simplified to [24-26]:

\[
\begin{align*}
A_{j,2k}, A_{j,2k+1} & = \text{Split(A}_j) \\
 d_{j+1,k} & = A_{j,2k+1} - \rho(A_{j,2k}) \\
 a_{j+1,k} & = A_{j,2k} + \mu(d_{j+1,k})
\end{align*}
\]

With this implementation of the forward Wavelet transform, \(z(x)\) has been driven to subsets, \(d_{j,k}\) and \(a_{j,k}\), which records high and low frequency events at the scale \(2^{-(j+1)}\) of \(z(x)\). The whole decomposition of \(z(x)\) is a simple repetitive scheme through rows and columns and all computations are carried out in-place. After the \(j^{th}\) step of decomposition in the scalar domain, an original surface signal \(z(x)\) is replaced with the wavelet series \(d_{1,}, \cdots, d_{j}, a_{j}\). It can be expressed by [27]:

\[
W_j[z(x)] = (a_{1,k}, d_{1,k}) = (a_{2,}, d_{2,}, d_{1,}) \\
= (a_{j,}, d_{j,}, d_{j-1,}, \cdots, d_{1,})
\]
The above equation is Mallat’s format arrangement of the wavelet transform coefficients. In practice, due to the lifting scheme’s in-place calculation property, the coefficients are organised in-place. Let input data \( z(i) \) with length \( i = 0...N-1 \), the output of the transform will also has length \( N \), then \( a_j \) is the subset with index \( n \cdot 2^j \), where \( n = (0,...,\frac{N}{2^j} - 1) \); \( d_j \) is the subset with index \( n \cdot 2^j + 2^{j-1} \), where \( n = (0,...,\frac{N}{2^j} - 1) \).

(a) the diagram of the lifting scheme

(b) the scheme of lifting Lazy wavelet

Fig. 4. Wavelet transform using the Lifting scheme

Figure 5 illustrates this schedule where \( j \) is the decomposed level of Wavelet transform in scalar domain [27].

3.4 Separation and extraction of frequency components of a surface

If a surface, \( z(x, y) \) is assumed to consist of a series of superimposed frequency components:

\[
z(x, y) = \eta(x, y) + \eta'(x, y) + \eta''(x, y)
\]  

(28)
Wavelet Analysis for the Extraction of Morphological Features for Orthopaedic Bearing Surfaces

Fig. 5. Wavelet transform

These surface components can be separated in the scalar domain by band pass filters. The transmission bands are based on different cut-off wavelengths for different frequency components. The wavelet coefficients, $d_i, \ldots, d_j$, can be considered outputs of a high frequency band, $1/\lambda \sim 1/\lambda$, and referred to as the roughness component, $\eta(x, y)$. $1/\lambda$ indicates the high frequency limited by the sampling interval, and $1/\lambda$ is the roughness frequency limitation. The $d_{j+1}, \ldots, d_j$ represents the output (waviness, $\eta''(x, y)$) of a sub-lowpass filter band $1/\lambda \sim 1/\lambda_\omega$. $1/\lambda_\omega$ is the waviness frequency limitation. $a_{\omega}$ are scalar coefficients that represent the output (form error, $\eta''(x, y)$) of a lowpass filter band, $1/\lambda_\omega \sim 1/l$; where $l$ is the sample length when $l=l_i$ or $l_j$. Using an inverse Wavelet transform, these surfaces can be recovered flexibly and immediately in the different transmission bands in terms of functional analysis requirements. The inverse Wavelet lifting Transform is performed simply by reversing the order of the operations and toggling negative to positive for all operations [27]. The relationship between the cutoff frequency and the wavelet scale is shown in table 1.

$$ $$

Where, $J = \text{int}(\log_2(\lambda / \Delta x) - 1)$, $J_{w} = \text{int}(\log_2(\lambda_{\omega} / \Delta x) - 1)$. Table 1 gives the relationship between the decomposition level and the cutoff wavelength.

In order to obtain $\eta(x, y)$ and $\eta''(x, y)$, the appropriate scalar coefficients, $a_{\omega}$, are set to zero and then the inverse transform is applied, and vice versa for $\eta''(x, y)$. In a similar way, a flexible reference surface can be obtained:

$$m(x, y) = z(x, y) - \eta(x, y) = z(x, y) - IW(d_i, \ldots, d_j)$$

where $J$ is a flexible selected decomposition level according to the cutoff wavelength of the reference surface. If a functional evaluation of surfaces is needed to cover all of the
topographical information from roughness, through multi-scalar events, to waviness, a functional surface can be built by:

\[
\eta(x,y) + \eta'(x,y) = \mathcal{IW}(d_i, \cdots, d_{jw})
\]

(31)

<table>
<thead>
<tr>
<th>Cut-off wavelength</th>
<th>The decomposition level of wavelet filter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3 (2^*3=16)</td>
</tr>
<tr>
<td>4</td>
<td>0.064 (0.08)</td>
</tr>
<tr>
<td>8</td>
<td>0.128</td>
</tr>
<tr>
<td>12</td>
<td>0.192</td>
</tr>
<tr>
<td>16</td>
<td>0.256 (0.25)</td>
</tr>
<tr>
<td>20</td>
<td>0.320</td>
</tr>
<tr>
<td>24</td>
<td>0.384</td>
</tr>
</tbody>
</table>

Table 1. The relationship between the level of the wavelet decomposition and sampling interval

The identification of multi-scalar events in the roughness and waviness bands is important in order to study the functional performance of the areal surface topography. These multi-scalar topographical events, \( \xi(x,y) \), such as peaks/pits and rings/valleys, hide in the roughness and waviness bands, and their wavelengths can cover a wide frequency range \( \frac{1}{\lambda_i} \sim \frac{1}{\lambda_{w,\lambda}} \). Wavelet coefficient sets over the transmission bands “naturally” record the information concerning their amplitude and location. These events can be easily captured by using an amplitude threshold, \( T_j \), to pick out the roughness and waviness. \( T_j \) is the value of an intersection of the probability curve of the cumulative amplitude distribution of each wavelet coefficient set. This process is based on an assessment that the amplitude distribution of each wavelet coefficient set, \( d_{j,k} \), belonging to the roughness and waviness components, would obey a normal distribution. If the absolute value of the amplitude is equal to or larger than \( T_j \), a thresholding estimator is applied.

\[
d_{j,k} = \begin{cases} 
0 & |d_{j,k}| < T_j \\
d_{j,k} & |d_{j,k}| \geq T_j 
\end{cases}
\]

(32)

Where, \( j \) is the decomposition level, \( k = n \cdot 2^i + 2^{i-1} \) with \( n = (0, \ldots, \frac{N}{2^i} - 1) \) are the indices of the coefficients. The absolute value of the peak amplitude is smaller than \( T_j \), the coefficient should be replaced by a zero. In the case of a detail coefficient being larger than or equal to \( T_j \), the coefficient should be retained. As a consequence, detail coefficients that represent only the information of the topographical events are obtained. From the experiments carried out, the threshold approaches the standard deviation of each wavelet coefficient set. The multi-scalar topographic features can then be built by using the 2D inverse discrete wavelet transform:

\[
\xi(x,y) = \mathcal{IW}(d_i', \cdots, d_{jw}')
\]

(33)

Where \( d' \) represent the thresholded wavelet coefficients.
3.5 Case studies on the bioengineering surfaces

Typical examples of much finer bioengineering surfaces from hip joint systems have been selected to show the behaviour of the spline wavelet filtering. During the Total Hip Replacement, a prosthetic head as shown in Figure 6a will replace the head of a bad or dead human femur. For the femoral heads of the hip joint system, it needs not only to support most of the body weight, but also to resist long term wear in service. The wear property of the femoral head surface is crucial for the lifetime of the whole hip joint system. The topographies of the femoral heads of the hip joint system were obtained by using a phase-shifting interferometer with a sampling area of $228 \times 203 \mu m$. The illustrations in Figure 7 show different topographies in polar and equatorial regions of worn heads.

Fig. 6. A typical artificial hip joint

Fig. 7. The typical bearing surfaces of worn metallic femoral heads
Fig. 8. Rough surfaces of worn metallic femoral heads with cut-off length $\lambda = 0.05\text{mm}$

Fig. 9. Waviness surfaces of worn metallic femoral heads with cut-off length $\lambda = 0.10\text{mm}$

Fig. 10. Form error surfaces of worn metallic femoral heads

Whatever region the surface is measured, it is part of a ball, and so contains a very smooth spherical form. It also has two different kinds of scratches: regular and shallow scratches, possibly produced by the manufacturing processing; and random deeper scratches,
resulting from functional performance in service. The latter scratches have a wider frequency band and higher amplitude, some with arc structures.

Using only one operation of wavelet filtering, the surface contents, roughness, waviness and form error can be detected and recognised. The outcomes of this are that rough, wavy and corresponding form error surfaces can be immediately separated and perfectly reconstructed within a flexible transmission bank. For instance, in order to indicate this transmission flexibility, the cut-off wavelengths of roughness may be selected as \( \lambda = 0.05 \, \text{mm} \). The cut-off length of waviness is limited by practical applications, in the above examples, \( \lambda = 0.10 \, \text{mm} \) is for the worn heads. From Figures 7-10, it can be seen that the accuracy of the surfaces of these components cover the levels from micrometers to nanometers. There are no distortions caused by peaks/pits and scratches and boundaries, which can often be found in normal surface filtering.

The identification of the morphological features of surface topography of orthopaedic joint prostheses has been carried out in light of the Wavelet model outlined above. In the first set

![Fig. 11. The measured and morphological surfaces of a worn ceramic head](image1)

![Fig. 12. The measured and morphological surfaces of a worn metallic head](image2)
of specimens, measurement datums of ceramic and metallic femoral heads were originally obtained by using a phase-shifting interferometer, with a 20X lens. The datum of the Diamond Like Carbon (DLC) specimen was obtained by employing a vertical-scanning interferometer using a 10X lens.

Figures 11-13(a) show lapped topographies in both the worn ceramic and metallic femoral heads with sampling area $300 \times 240 \mu m$. The worn ceramic surfaces look smooth and have some deep scratches. The worn metallic surfaces have two different kinds of scratches, regular shallow scratches, possibly produced by manufacturing processing and the random deeper scratches, resulting from functional performance in service. The latter scratches have a wider frequency band and higher amplitude, some with arc structures. The following Figure 13a is DLC surface with a morphological structure consisting of relatively large pits.

Figures 11-13(b) show how the Wavelet model has removed roughness, waviness and form deviation, revealing morphological surfaces. As illustrated, the morphological features are the dominant factors of the bearing surface structure of the new or wear heads, and roughness and waviness do not seem to heavily influence the functional performance of the head in service due to their relatively low levels. Examining these different surfaces (Figure 11-13), it can be seen that: (1) the components of waveforms of traces resemble each other with no relative phase shift in the sampling area; (2) the morphological information on the 3D bearing surfaces, conveyed by the wavelet model, is recorded completely and (3) with no running-in and running-out lengths. These morphological surfaces therefore have excellent refinement accuracy, which is suited to the need for assessment of a range of functional properties of the components and study of functional performance of bearing surfaces. The information and revealing the presence of these outliers, provided by Wavelet analysis can then be fed back to monitor manufacturing processes or study tribology and wear; or to study actual contact stress, loaded area and asperity volume, and additionally lubrication regimes occurring during the initial stages of wear.

4. The third generation complex wavelet model for surface feature extraction

The extraction of morphological features such as linear/curved scratches and plateaux with direction/objective properties require more consideration of how wavelets are used, for
example: (1) **Automotive**. Plateau honed surfaces are produced by two machining processes. The final surface always includes the rough machining information (deep valleys). (2) **Biomedical engineering**. Surface topographies in different material heads are normally combined with morphological features from different wear stage in service, such as regular shallow scratches, random deeper scratches, arc scratches.

Morphological features in surface texture are not always of a large-amplitude and isolated. In many cases, morphological features have direction properties; they mix with the harmonic roughness and waviness and have similar amplitude scales and frequency bands. The first and second generation wavelet methods for extraction of such events in the surface texture is less effective at extracting these features. The reason is that the first and second generation wavelet models are used in there maximally decimated real discrete wavelet transform form, and lack the shift invariance property. This means that small shifts of surface signal (step height, linear and curve-like features) can cause large variations in the distribution of energy between real wavelet transform coefficients at different scales. The directional morphological features from these wavelet techniques are also destroyed by edge artifacts. Consequently, under the first two generation wavelet models, the accurate extraction of direction/objective morphological features is not efficient.

A new method for extraction of direction/objective morphological features of surfaces has been proposed using a biorthogonal dual tree complex wavelet transform (DTCWT) [30,31]. It attempts to give affine invariance, with independence of the reference frame for the measurements, and also perfect reconstruction, limit redundancy and have efficient computation.

### 4.1 Theory of DT-CWT

For the extraction of direction/objective morphological features of surfaces, the DT-CWT attempts to give affine invariance, with independence from the reference frame for the measurements, and also perfect reconstruction, limit redundancy and efficient computation.

The complex wavelet model is based on the Z-transform theory of linear time invariant sampled systems. The main approach of DT-CWT operation is to use two-channel filter banks in the real and imaginary parts as shown in Figure 14. A Q-shift filter design technique is used to construct the lowpass FIR filter to satisfy a linear-phase and perfect reconstruction condition, as is demonstrated below [28-29]:

\[
\begin{cases}
H_{1+}(z) = H_{1}(z^2) + z^{-1}H_{1}(z^{-2}) \\
H_{1}(z)H_{1}(z^{-1}) + H_{1}(-z^2)H_{1}(-z) = 2
\end{cases}
\]  

The one-dimensional complex wavelet decomposes a signal \( f(x) \) in terms of a complex shifted and dilated mother wavelet:

\[
f(t) = \sum_{a} a_j\phi_j(x) + \sum_{b} d_j\psi_j(x)
\]  

where \( \phi_j = \phi_{j} + \sqrt{-1}\psi_{j} \) and \( \psi_j = \psi_{j} + \sqrt{-1}\phi_{j} \). \( J \) denotes the coarsest scale can be calculated by:

\( J = \text{int}(\log_{2}(\lambda_c / \Delta x) - 1) \), where \( \lambda_c \) is the cut-off wavelength, and \( \Delta x \) is the sampling spacing.

The \( \psi_j \) and \( \phi_j \) are real wavelets. \( a_j \) and \( d_j \) refer to the smooth coefficients at level \( J \)
and the detail components of coefficients at the level \( j \), respectively. The complex wavelet transform is essentially a combination of two different real wavelet transforms, in one dimension, the \( \{ \phi_{j,i}, \psi_{j,i}, \phi'_{j,i}, \psi'_{j,i} \} \) can be interpreted as a wavelet tight frame with a redundancy factor of two. The real and imaginary parts of the complex wavelet transform are computed using separate filter bank structures in different trees.

Fig. 14. Framework of the dual tree complex wavelet transform (DT-CWT)

For two dimensions, the decomposition of complex wavelet transform can be represented by the dilations and translations of a complex scaling function and six complex wavelet functions:

\[
f(s_1, s_2) = \sum_{i \in \mathbb{Z}^2} a_{i,j} \phi_{0,0} (s_1, s_2) + \sum_{b \in B} \sum_{i \in \mathbb{Z}^2} d_{j,i}^b \psi_{j,i}^b (s_1, s_2)
\]  

(36)

where, \( B = \{ \pm15, \pm45, \pm75 \} \) for six subbands’ directions. For the separable two dimensional complex wavelet transform based on one dimensional complex \( \phi, \psi \), we have [28-30]:

\[
\begin{align*}
\psi_{00}^{+} &= \phi(s_1) \psi(s_2), \quad \psi_{00}^{-} = \phi(s_1) \overline{\psi}(s_2) \\
\psi_{45}^{0} &= \psi(s_1) \psi(s_2), \quad \psi_{45}^{-} = \psi(s_1) \overline{\psi}(s_2) \\
\psi_{75}^{0} &= \phi(s_1) \phi(s_2), \quad \psi_{75}^{-} = \phi(s_1) \overline{\phi}(s_2)
\end{align*}
\]

(37)

where \( \overline{\phi}, \overline{\psi} \) is the complex conjugate of \( \phi, \psi \). Similar to the multi-resolution analysis of real wavelets, the above decomposition can be summarized as a decomposition of \( L^2 \) space into a series of subspaces \( \{ W_j, W'_j, \ldots, W_j, V_j \} \), here \( W_j \) and \( V_j \) respectively denote the wavelet bandpass subspace and the scale lowpass subspace at level \( j \). They satisfy [28-30]:

\[
V_{j+1} \subset V_j = V_{j+1} \oplus W_j, \quad W_j \perp W_{j'}, j \neq j'
\]

(38)
4.2 Transmission characteristics of DT-CWT

To extract different frequency components such as the roughness, waviness, form and form error accurately, the surface filter requires a near ideal transmission characteristics which include the zero or linear phase transmission and a ‘steep transmission curve’ amplitude transmission.

4.2.1 Phase transmission characteristics

A zero phase filter tends to preserve the shape of the surface components in the pass band region of the filter. This is very important for applications, as surface topographies are often composed of a wide range of frequency components. Zero or linear phase transmission ensures that the filtering results have no phase distortion and ensures the proper registration of different frequency components that make up these features. A digital filter has zero phase when its frequency response \( H(\omega) \) is a real function so that \( H(\omega) = H'(\omega) \). From the symmetry property of the Fourier transform and considering that only real impulse response \( h(n) \) is applied to real input to obtain real output, the necessary and sufficient condition for a zero phase filter is: \( h(n) = h(-n) \).

![Fig. 15. DT-CWT impulse response with 2 levels of decomposition](image)

(a) Odd position impulse response (b) Even position impulse response

The difference between zero phase and linear phase property of filter lies only in the shift of the co-ordinate system origin which has no practical significance. From Eq.8 we know that to get zero or linear phase property, the filter impulse response or the filter weighting function should be symmetric. Figure 15 shows the impulse response of a DT-CWT low pass and high pass filter. From the figures it is clear that both the odd position and the even position impulse response have symmetrical shape, and the odd and even position impulse responses are very similar. For the real DWT filter, the odd and even position impulse responses are different. Therefore, the DT-CWT can produce the shift-invariance property with zero/linear phase, whereas the DWT can not produce the shift-invariance property [31].

4.2.2 Amplitude transmission characteristics

For surface analysis, filters need a ‘steep transmission curve’, which requires that the low-pass filter with a steep cut-off such that a negligible amount of higher frequency information
passes and vice versa for high-pass filter. The “steep transmission curve” property is good for separation of different frequency components with a sharp attenuation at the cut-off wavelength, which leads to more precise separation of roughness, waviness and form components with no aliasing, and leads to the good feature location property. The Gaussian filter is designed to have 50% transmission at the cutoff frequency. The high pass filter and low pass filter have complementary relations, which allow the calculation of waviness by simply subtracting the roughness from the form removed raw profile. The DT-CWT also has very good amplitude transmission characteristics compared to the Gaussian filter. Figure 16 shows the highpass and lowpass amplitude transmission characteristics and the subband transmission characteristics of the first four decomposition levels. Figure 16(a) is the inverse transform of the level one DT-CWT transform of an impulse signal, and shows that the highpass and lowpass filter are intersected at the normalised frequency (frequency divided by the maximum frequency so it is normalised to the range (0,1)) 0.5 with 50% amplitude transmission. Accordingly, for an original input bandpass signal, similar to the Gaussian filter, 50% transmission at the cutoff frequency allows the calculation of waviness by simply subtracting the roughness from the raw profile. The DT-CWT has the “steep transmission curve” property so that it can separate out different frequency components effectively. Figure 16(b) shows the transmission characteristics of each of the four subbands levels from the DT-CWT decomposition. This shows that every adjacent subband has 50% transmission rate at the intersection frequency [31].

![Image](attachment:image1.png)  
(a) Transmission characteristics  
(b) Subband transmission characteristics  

Fig. 16. Transmission characteristics of a DT-CWT (using normalized frequency)

Figure 17 compares the transmission characteristics of the DT-CWT and ISO Standard Gaussian filter using the same cutoff frequency. The continuous lines are the results from the DT-CWT and the dashed lines are the results from the Gaussian filter. Although there exists some very little and negligible ripples in the pass band, it is clear that the DT-CWT always have a steeper transmission curve than the Gaussian filter near the cutoff frequency. Therefore the DT-CWT has a better frequency resolution than the Gaussian filter. Considering both the zero/linear phase property and the ‘steep transmission curve’ of the amplitude transmission characteristic, the DT-CWTs are suitable for the separation of different surface topographical features. Besides the separation of linear/curve-like features, the DT-CWT is also very good for the separation of roughness, waviness, form and form error such surface frequency components.
Wavelet Analysis for the Extraction of Morphological Features for Orthopaedic Bearing Surfaces

(a) normalized cutoff frequency 0.5
(b) normalized cutoff frequency 0.25
(c) normalized cutoff frequency 0.125
(d) normalized cutoff frequency 0.0625

Fig. 17. Comparison of the high pass and low pass transmission characteristics of the CWT and the Gaussian filter (continuous line: DT-CWT filter; dashed line: Gaussian filter)

### 4.2.3 shift-invariant property

A linear feature and a circular feature have been simulated, as shown in Figure 18 and Figure 19. Figure 18(a) is a simulated 3D groove feature that is neither parallel nor vertical to the measurement direction. Figure 18(b) and (c) shows the results processed by a complex wavelet transform and a real wavelet transform respectively. From left to right in Figure 4.2, the decomposed components of the input signal are reconstructed by the complex wavelet coefficients at scalar level 1, 2 and 3 and the scaling function coefficients at scalar level 3 correspondingly. From the results, it can be seen that in each scale, the extracted feature from the input signal using the complex wavelet has a very smooth edge along the groove and the same amplitude energy distribution in every position. Comparison with Figure 18(b), the extracted features using the real wavelet as shown in Figure 18(c) have a significant shift aliasing, with irregular edges along the groove feature [28-31].

From new the wavelet model shown in the above section, it can be seen that the dual tree complex wavelet offers six bandpass components of complex coefficients at each level, which are strongly oriented at angles of ±15°, ±45° and ±45°. The special point of this kind of complex wavelet is that it provides a true directional selectivity for curve-like features. Figure 19(a) is a simulated 3D circular valley feature, and Figure 19(b) and (c) 6 shows the outcomes as reconstructed from complex wavelet and real wavelet coefficients respectively at scalar levels 1, 2 and 3.
Fig. 18. Decomposition by using the real wavelet (UP: DTCWT, Down: Real DWT)

(a) A simulated groove feature

(b) Wavelets (level 1) Wavelets (level 2) Wavelets (level 3) Scaling (level 3)

(c) Wavelets (level 1) Wavelets (level 2) Wavelets (level 3) Scaling (level 3)

Fig. 19. Decomposition by using the real wavelet (UP: DTCWT, Down: Real DWT)

(b) Wavelets (level 1) Wavelets (level 2) Wavelets (level 3) Scaling (level 3)

(c) Wavelets (level 1) Wavelets (level 2) Wavelets (level 3) Scaling (level 3)

Fig. 4.4 A simulated 3D circular feature

(b) Wavelets (level 1) Wavelets (level 2) Wavelets (level 3) Scaling (level 3)

(c) Wavelets (level 1) Wavelets (level 2) Wavelets (level 3) Scaling (level 3)

Fig. 19. Decomposition by using the real wavelet (UP: DTCWT, Down: Real DWT)
4.3 Case studies
A group of typical examples of surfaces from engineering manufacture have been selected to show the metrological characteristics of the DT-CWT.

Figure 20 shows a typical peripheral milling surface profile filtered by the DT-CWT and the Gaussian filter with different cutoff wavelengths (0.64mm, 2.56mm). Both the DT-CWT and the Gaussian filter can extract the roughness mean line very well with no phase distortion. But the DT-CWT filtering results are smoother than the Gaussian filtering results, which shows that the DT-CWT filter has a steeper amplitude transmission characteristic than the Gaussian filter. Thus the DT-CWT has a better ability to separate different frequency components. Figure 21 and Figure 22 are grinding and turning surface profiles respectively. From the filtering results they also illustrate the good transmission characteristics including phase and amplitude properties of the DT-CWT.

The metrological characteristics of the DT-CWT investigated above can also be extended to areal surface analysis. Figure 23 is a worn hip joint prosthesis surface, the measured raw...
data includes waviness and form/form error that was introduced by designed shape and measurement error. By using DT-CWT the form, waviness and roughness have been successfully extracted and separated. It is also very clear that the roughness and waviness surfaces have no distortion.

![Graphs showing raw data, extracted form, waviness, and roughness](image)

Fig. 22. Turning surface profile filtered by DT-CWT and Gaussian filter

![3D images showing measured raw data, extracted form, waviness, and roughness](image)

Fig. 23. Separation of the different surface components
The above research has shown that shift-invariance is a very important property. In general, the shift-invariance is independent of the reference frame for a surface measurement. In particular it is important for many approximation and filtering methods.

To verify the special properties of the third generation wavelet model, a group of typical examples of engineering and bioengineering surfaces are shown in Figures 4.7a ~ 4.10a, in which some significant features have been selected. These will be used to demonstrate the performance of the dual tree complex wavelet transform.

The first example shows two measured surface from different positions of the same worn hip-joint. Figure 24(a) (b) are raw measured surfaces, including the form, form deviation, waviness and roughness components. Figure 24(c) and (d) are the reconstructed feature surface from (a) and (b) respectively. The deep valleys are reconstructed using complex wavelets in which not only the form and some harmonic components are removed very effectively, but also the shape of the scratches are retained very well. There is no affine aliasing, no new artifacts and the edge of the deep valleys are preserved perfectly. The filtering results help to evaluate the worn rate of the hip joint at different position and further can help to improve the design.

(a) slightly worn  (b) heavily worn

(c) feature surface of  (d) feature surface of

Fig. 24. Cylinder surface analysis using the complex wavelet
The following examples in Figure 25(a) to Figure 27(a) are a series of surfaces from different material femoral heads. Figure 25(a) shows a lapped topography surface in a worn metallic femoral heads. The worn metallic surfaces have two different kinds of scratches, regular shallow scratches, possibly produced by manufacturing processing and the random deeper scratches, resulting from functional performance in service. The latter scratches have a wider frequency band and higher amplitude, some with arc structures. Figure 26(a) is a surface topography from a diamond-like-carbon head with a morphological structure consisting of relatively large pits and a deep scratch. Figure 27(a) is a new ceramic surface which looks smooth and has some shorter and deeper scratches.

Looking at these surfaces, it must be noted that the most important factors that affect the life of the femoral head are the isolated peaks/pits and scratches rather than the nano-scalar roughness. From the function assessment point of view, the isolated peaks/pits and scratches will significantly interfere with the wear mechanics and tribological properties[6-9,27].

Using the complex wavelet, all the isolated features are extracted precisely as shown in Figure 25(b) to Figure 27(b), and the shape of the features are reconstructed very well and all the features’ edges are preserve perfectly.

![Fig. 25. Worn metallic surface analysis using the complex wavelet](image1)

![Fig. 26. Diamond like carbon surface analysis using the complex wavelet](image2)
Wavelet Analysis for the Extraction of Morphological Features for Orthopaedic Bearing Surfaces

5. Complex ridgelet transform

Wavelets have the good performance for piecewise smooth functions in one dimension, i.e., fully efficient at representing point-like isolated morphological features. Unfortunately, such is not the case for higher dimensional features because wavelets ignore the geometric properties of objects. So the function of wavelet techniques for extracting line/curve features is limited.

To extend the directional sensitivity for wavelet analysis, an approach based on a combination of the wavelet and Radon transform is used. This was initially developed for edge detection and compression in the field of signal processing.

Ridgelets were introduced by Candès and Donoho to deal effectively with line singularities by mapping a line singularity into a point singularity through the Radon transform, and then using the wavelet transform on each projection in the Radon transform domain [32-34]. The weakness of the traditional ridgelets transform is the lack of shift invariance due to the real DWT used.

5.1 Complex ridgelet transform

A bivariate complex ridgelet \( \psi_{a,\theta,b} \) in \( \mathbb{R}^2 \) can be defined as [32-37]:

\[
\psi_{a,\theta,b}(x) = a^{-\frac{3}{2}} \psi' \left( \left( x_1 \cos \theta + x_2 \sin \theta - b \right) / a \right)
\]

Here, \( a > 0 \) is a scale parameter, \( \theta \) is an orientation parameter, and \( b \) is a location scalar parameter. This function is constant along lines \( x_1 \cos \theta + x_2 \sin \theta = \text{const} \), while its transverse is a complex wavelet \( \psi' = \psi' + \sqrt{-1} \psi' \). Here, \( \psi' \) and \( \psi' \) are themselves real wavelets. If the real and imaginary part of complex wavelet can be viewed as two ‘fat’ points, then the complex ridgelet can be interpreted as two ‘fat’ lines so that it is specially adaptive to analysis the line ridge/valleys of surface topography. Figure 28 shows a real ridgelet function.

The continuous Complex Ridgelet Transform (CRIT) for an integrable bivariate function \( f(x) \in L^1(\mathbb{R}^2) \) is defined as:

\[
\text{www.intechopen.com}
\]
The reconstruction formula is given as:

\[ R_f(a,b,\theta) = \int \psi_{a,b,\theta}^*(x) f(x) dx \]  
\[ (40) \]

Point and line singularities are related by the Radon transform. By applying the 1-D DT-CWT on the projections of the Radon transform the complex ridgelet transform can be rewritten as:

\[ R_f(a,b,\theta) = \int R_r(\theta,t) a^{-\frac{1}{2}} \psi^*\left((t-b)/a\right) dt \]  
\[ (42) \]

The Radon transform is denoted as:

\[ R_r(\theta,t) = \int f(x_1,x_2) \delta(x_1 \cos \theta + x_2 \sin \theta - t) dx_1 dx_2 \]  
\[ (43) \]

Where, \( \delta \) is the Dirac delta function.

Fig. 28. A typical ridgelet function \( \psi_{a,b,\theta}(x,y) \)

### 5.2 Digital complex ridgelet transform

From above equations, one can see that the basic strategy for computing the CRIT is to first calculate the Radon transform \( R_r(\theta,t) \), then to calculate the 1D-CWT of the projections \( R_r(\theta, \cdot) \). For the calculation of the Radon transform, numerous digital methods have been devised. However, most of them were not designed to be invertible transforms for digital surfaces or images. Alternatively, the finite Radon transform (FRAT) theory provided an interesting solution for finite length signals. According to the practical requirements of surface characterisation, we use the digital form of the CRIT based on the FRAT and DT-CWT.
5.2.1 Finite Radon Transform [34,38-39]

The FRAT is defined as summations of surface pixels over a certain set of lines. These lines are defined by a finite geometry in a similar way as the lines for the continuous Radon transform in Euclidean geometry. Denote the group $G = \mathbb{Z}^2_p$ to be the Cartesian product $\mathbb{Z}_p \times \mathbb{Z}_q$ of two exemplars from the cyclic group $\mathbb{Z}_p = \{0, 1, \cdots, p-1\}$ with addition modulo $p$, where $p$ is a prime number, and let $N = \{0, 1, \cdots, p\}$.

This group has $p+1$ nontrivial subgroups

$$H_j = \{(k,l) \in G; \; li = k \pmod{p}\}, \quad 0 \leq i < p$$

$$H_j = \{(k,0) \in G; \; k \in \mathbb{Z}_p\},$$

(44)

The cosets of the factor group $G/H_j$ are indexed by $j \in \mathbb{Z}_p$ in the following way

$$H_j^i = \{(k,l) \in G; \; li + j = k \pmod{p}\}, \quad 0 \leq i < p$$

$$H_j^i = \{(k,j) \in G; \; k \in \mathbb{Z}_p\}.$$  

(45)

The Radon projection of a function $f$ on $G$ is given by

$$\Lambda f (H_j^i) = \frac{1}{p} \sum_{i=0}^{p-1} \sum_{l=0}^{p-1} f(k,l) \delta_i (\pi_i (k,l)), \quad i \in N, j \in \mathbb{Z}_p$$

Here, the function $\pi_i$ is a variance of the factor mapping of $G$ on $G/H_j$. We have

$$\pi_i (k,l) = k - li \pmod{p}, \quad 0 \leq i < p$$

$$\pi_i (k,l) = l \pmod{p}$$

(47)

For surface topography or for an image, the coset $H_j^i$ denotes the set of points that make up a line on the lattice $G$. Particularly $H_0^j$ and $H_p^j$ denote the horizontal and vertical lines respectively. $\pi_i$ denotes the set of lines that go through a point $(k,l) \in G$. As in the Euclidean geometry, the line $H_j^i$ on the affine plane $G$ is uniquely represented by its slope $i \in N$ and its intercept $j \in \mathbb{Z}_p$. $H$ has $p^2$ points, $p^2 + p$ lines, every point $(k,l) \in G$ lies on $p+1$ lines, every line contains $p$ points. Moreover, any two distinct points on $G$ lie on just one line. For any given slope $i \in N$, there are $p$ parallel lines to provide a complete cover of the plane $G$.

From the finite geometry property, for zero-average functions, the forward and inverse formula can be written as:

$$F = \Lambda f = \frac{1}{p} \sum_{(i,j) \in \mathbb{Z}_p^2} f(k,l) = \frac{1}{p} \sum_{(i,j) \in G} f(k,l) \delta_{ij} (k,l) = \left\langle f, \frac{1}{\sqrt{p}} \delta_{ij} \right\rangle$$

$$f = \mathcal{F} F = \frac{1}{p} \sum_{(i,j) \in \mathbb{Z}_p^2} F(i,j) = \frac{1}{p} \sum_{(i,j) \in G} \sum_{(k,l) \in \mathbb{Z}_p^2} f(k',l')$$

$$= \frac{1}{p} \sum_{(k,l) \in G} f(k',l') + f(k,l) = f(k,l)$$

(48)
Where $\delta_{ij} : G \rightarrow R$ is the Dirac delta function, defined as:

$$\delta_{ij}(k,l) = \begin{cases} 1 & (k,l) \in H_i^j \\ 0 & \text{elsewhere} \end{cases}$$

The inverse FRAT algorithm has the same structure and is symmetric with the algorithm of the forward FRAT.

### 5.2.2 Digital Ridgelet Transform

The complex wavelet basis is defined as:

$$\{w^{(i)}_m, \ m \in Z_p, \ i \in N\} \quad (49)$$

The digital FRIT can be integrated as:

$$CFRIT_i(m,n) = \left\langle A, f, w^{(i)}_m \right\rangle = \sum_{m} w^{(i)}_m(j) \left( f, \frac{1}{\sqrt{p}} \delta_{i} \right) = f, \frac{1}{\sqrt{p}} \sum_{m} w^{(i)}_m(j) \delta_{i}$$

Therefore the basis functions of the discrete complex ridgelet transform can be written as

$$\rho_{i,m} = \frac{1}{\sqrt{p}} \sum_{m} w^{(i)}_m(j) \delta_{i} \quad (51)$$

Here, although the $\left\{ \frac{1}{\sqrt{p}} \delta_{i} \right\}$ is not an orthonormal system, if we take the $p+1$ orthonormal bases for $L^1(Z_p)$, $\{w^{(i)}_m, \ m \in Z_p\}$ with $w^{(i)}_m = \text{const}$, the system $\{\rho_{i,m} : i = 0, \ldots, p; m = 1, \ldots, p-1\} \cup \{\rho_0\}$ is an orthonormal base for $L^1(Z_p)$ where $\rho_0(k,l) = 1/p, \forall (k,l) \in G$.

Figure 29 gives the basic procedure when using the proposed CFRIT to analysis surface.

### 5.3 Case studies

Figure 30 shows 16 shifted versions of the image (at the top) and their subspace reconstructed components in turn from the coefficients at levels $j \leq j_0 = 4$ using the CFRIT (left) and real FRIT (right). In order to see the effects clearly, only the centre of the profiles of these images is shown. Each shift is displaced down a little to give a waterfall style display. The output of CFRIT is the modulus of the complex coefficients. Note that summing these components the input image can be reconstructed perfectly. Good shift invariance is seen from the fact that the shape and amplitude of each of the reconstructed components by CFRIT hardly varies as the input is shifted. In contrast, the reconstructed components using FRIT vary considerably with each shift.

Our next test shows the good performance of CFRIT for denoising an image with line singularities. We consider an artificial image with a deep scratch that is contaminated by an additive zero-mean Gaussian white noise of variance $\sigma^2$. The denoising includes the following steps: (1) Transform the noisy image using CFRIT; (2) Hard-thresholding of the
Wavelet Analysis for the Extraction of Morphological Features for Orthopaedic Bearing Surfaces

Fig. 29. Directional feature extraction

Fig. 30. The reconstructed components (centre profiles) at levels 1 to 4 of 16 shifted image with a stepped edge using the CFRIT (left) and real FRIT (right). Each shift is displaced down a little to give a waterfall style of display.
Fig. 31. Denoising an image with line singularities using the DWT (upper right), DT-CWT (lower left), and CFRIT (lower right). The upper left image is a noisy image.

Fig. 32. Feature extraction of hone surface using the DWT (upper right), FRIT (lower left), and CFRIT (lower right). The upper left image is the original data (form removed).
coefficients using the universal threshold $T = \sigma \sqrt{2 \log N}$ (where $N$ is the number of pixels).

(3) Reconstructing the thresholded coefficients. For comparison, the same uniform threshold value is also applied to the DWT and DT-CWT based algorithms.

It can be seen from Figure 31 that the CFRIT is effective in recovering straight edges, as well as in terms of the signal to noise ratio (SNR). The CFRIT reconstruction does not contain the undesirable artefacts along edge that one finds in the wavelet reconstruction. The simple thresholding scheme for CFRIT is effective in denoising the piecewise smooth image with line singularities. This is because the linear singularities are represented by a few significant coefficients in the CFRIT domain, whereas random noisy singularities are unlikely to produce the similar amplitude coefficients. The DT-CWT is relatively superior to the DWT in reducing the artefacts due to its shift invariance and good directional selectivity. From the view of the hybrid approach, the CFRIT just combines the multiresolution analysis of DT-CWT with the anisotropy of the Radon transform.

Figure 32 shows a honed surface from an engine cylinder. As is well known, the most important features that influence the performance of cylinders are the deep scratches, the distribution and amplitude will considerably influence the flow of gas or air in the pressure balance of an engine. It can be seen that the DWT still exhibits numerical embedded blemishes (i.e., pits/peaks) in the extracted honed surface. Setting higher thresholds to remove these would cause even more of the intrinsic linear scratches to be destroyed or missed. In addition, the non-smooth aliasing along the scratches is clearly visible. In the FRIT result, although the pits/peaks have been removed efficiently, the edges of the linear scratches are not very smooth with aliasing due to the lack of the shift invariance property. In the CFRIT result, not only are the peaks/pits in the honed surface removed effectively, but also the shapes of the extracted scratches are well retained. There is also no affine aliasing, and the edges of the deep valleys are preserved perfectly.

To compare the FRIT with the general wavelet thresholding for line and curve feature extraction, the above form removed surface was shown in Figure 33(a) as a grey topview image. The feature extraction results of the general wavelet thresholding, CFRIT are shown in Figures 33(b) and (c) respectively. Comparing the results of the general wavelet thresholding and the CFRIT, the features from the wavelet thresholding have blur margin and are discontinuous in some change points while those from CFRIT are smooth and continuous.
6. Conclusions

In this chapter, the early families of orthogonal wavelets and the three generations of biorthogonal wavelet transforms, developed later, for the extraction of morphological structures from micro/nano scalar surfaces in the field of bioengineering, have been introduced. The chapter’s aim is to create a “tool box” of wavelet techniques capable of complex analysis and interpretation of surface topography data; leading to the extraction of functionally critical morphological features from micro/nano scalar surfaces of orthopaedic joints for in-vitro and clinically retrieved applications.

The orthogonal wavelets have been used for analysis of multi-scalar surfaces in engineering in which the phase distortions were neglected. The main advantages of biorthogonal wavelets are that they are projections (into an associated wavelet scale space) with linear phase (leading to real outputs without frequency aliasing and phase distortion) and a traceable location property.

The first generation biorthogonal wavelet transform overcame the phase distortion disadvantage of orthogonal wavelet transform, thus were suitable for surface filtering. The first generation model was successfully applied in de-noising areal surface stylus instruments.

The second generation model was built using the lifting scheme, which does not use the Fourier transform as a prerequisite, and can be implemented in three easy stages namely, splitting, prediction and updating. The models for second generation wavelets have helped the Bio-engineering industry in successful separation, extraction and reconstruction of isolated morphological features. For example, large peaks/pits in hip joint heads. The accuracy of the bearing surfaces of these medical components is from micrometers to nanometers. The reconstruction of morphological features has helped friction, contact stress and load area (which are evident during the initial stages of wear) to be more accurately assessed.

The first/second generation wavelet models were built by the real discrete wavelet transform that is less effective in the extraction of certain morphological features, such as linear/curved scratches and plateaux with direction/objective properties in micro-structured surfaces, for example, plateau honed surface, steel and stainless steel sheets, bio-engineered surfaces. Real discrete wavelet transforms have a lack of shift-invariance and have less sensitivity to direction and anisotropy, as a result, they have good performance at representing sparse zero-dimensional or point singularities, but not higher dimensional objects such as linear/curved scratches.

The third generation wavelet models are based on novel mathematics – the Complex Wavelet Transform (CWT). It is then proposed to explore in detail two types of tools: 1) The DT-CWT for shift-invariant extraction of point-like morphological features; and 2) The Complex Finite Ridglet transform (CFRIT) for linear/curve morphological features.

7. Acknowledgments

The authors would like to thank the EPSRC UK for supporting this research work under its programme EP/F032242/1. The corresponding author X. Jiang gratefully acknowledges the Royal Society under a Wolfson-Royal Society Research Merit Award and the European Research Council for its ‘Ideal Specific programme’ ERC-2008-AdG 228117-Surfund.
8. References
