COMMUNICATIONS ON
PURE AND APPLIED ANALYSIS
Volume **xx**, Number **0x**, xxx **200x**

Website: http://AIMsciences.org

pp. **xx–xx**

THE RADON MEASURE FORMULATION FOR EDGE DETECTION USING ROTATIONAL WAVELETS

Liming Zhang, Tao Qian

Faculty of Science and Technology
University of Macau
Av. Padre Tomas Pereira, Taipa, Macao SAR China

Qingye Zeng

Department of Image Engineering
Chinese Academy of Sciences
Beijing 100101,China

Abstract. Based on a mathematical model involving Radon measure explicit computations on convolution integrals defining continuous (integral) wavelet transformations are carried out. The study shows that the truncated Morlet wavelet significantly depends on a rotation parameter and thus lay a foundation of edge detection in pattern recognition and image processing using rotational (directional) wavelets. Experiments and algorithms are developed based on the theory. The theory is further generalized to the n-dimensional cases and to a large class of rotational wavelets.

1. **Introduction.** Two main categories within wavelet theory are the continuous wavelet transform (CWT) on the one hand, and discrete orthogonal wavelet transform (DWT) developed through multi-resolution analysis on the other hand. They enjoy more or less opposite properties and both have their specific fields of applications. The DWT has proven to be a successful technique for, e.g., data compression, whereas the CWT is especially appropriate for signal analysis [1].

The one dimensional (**1D**) discrete wavelet transform and the simplest $2D$ discrete wavelet transform [2] are obtained by analyzing the original signal with dilated and translated versions of the wavelet function. The $2D$ continuous wavelet uses also rotated versions of the wavelet function for such analysis. The higher dimensional discrete wavelets in general also have the capacity to detect directional properties of images, but they are not through introducing a variable rotation as continuous wavelet [3]. The continuous wavelet with rotational effect is easier to use in practice. The rotation effect enables the $2D$ continuous wavelets to analyze the directional features of image signals.

The WT used in computer vision is a relatively new concept – about 10 years old – but yet there are quite a number of articles published involving use of this technique. Most of the papers are in the area of image compression and reconstruction. Continuous wavelets with a rotation parameter are proposed in pattern recognition in the literature [3-7]. What has been done along the direction may be roughly classified into three categories: (i) Certain characteristic properties, in non-quantitative terms in relation to the rotational Morlet wavelet function. Its Fourier transform and the corresponding wavelet transform were observed. One may predict from those observations the effectiveness of the rotational wavelets [3-5] in pattern recognition. (ii) Simple analysis based on some characteristic quantities of those analyzing wavelets, their Fourier

1991 *Mathematics Subject Classification.* 37C45.

Key words and phrases. rotational wavelets, edge detection, Morlet wavelet, Radon measure.

The work was supported by University of Macau Research Grant RG065/03-04S/QT/FST (joint by the first two authors) and Macao Science and Technology Development Fund 051/2005/A.

transforms (spatial frequency), and their reproducing kernels etc. were carried out. One deduces in those computations scale-resolving power, angle-resolving power and minimum grid to discretize the integrals related to the continuous wavelet transformations (CWT) and the inversion formulas [3]. (iii) Experiments and tests [3-5] were carried out. Comprehensive introductory knowledge and applications of rotational wavelets may be found in a series papers of J.-P. Antoine et al ([3,4,8,9]). On one side is the observations on the integrals defining wavelet transforms, and on the other side is the effectiveness in experiments of the use of the rotational wavelets. It seems to the authors, however, there exists a gap, for there is no mathematical proof for the effectiveness. This paper proposes a mathematical theory to fill in the gap. The theory also suggests how to use the method in practice.

Note that in this work only continuous (integral) wavelet transformations are concerned. Discrete wavelets are only for the comparison purpose. The terminology "rotational" in some literature is called "directional". The writing plan of the paper is as follows. Section 2 gives the basic mathematical notation and terminology adopted in the whole paper. They are standard and the reader may skip over the section and directly go to Section 3. Section 3 contains a methodology discussion on how one would possibly define the concept *edge*. Section 4 introduces the Radon measure analysis for edge detection. The section concentrates in the $2D$ case and the main result Theorem 4.1 is the foundation of the method. The proof of Theorem Th.4.1 is postponed to Section 6 where it is given for the general nD cases. Section 5 contains applications of the theory to edge detection. Section 6 extends the theory to the higher dimensional cases (nD cases). The section provides rigorous proofs that can be treated as an appendix for the readers who not only concern about applications but also mathematical reasoning. There we prefer to write the theory in the general nD setting. That is because, on one side, there is essentially no difference between $2D$ and nD cases and writing in the nD cases makes the mathematical idea clearer. On the other hand, applications for higher-than-two dimensional wavelets incorporated with rotational variables have been concerned in practice. Indeed, some physical phenomena are of a multi-scale nature. Higher dimensional ($3D$) pattern recognition using rotational wavelets may be found in image processing and fluid dynamics , $3D$ modelling of faces, the appearance of coherent structure in turbulent flows, or the disentangling of wave train in acoustics (See, for example, [3,4,8,9] and other relevant articles by Antoine et al) etc.

2. **Preliminaries.** Throughout the paper $\mathbf{Z}$ will denote the set of *positive integers*; $\mathbf{R}$ the set of *real numbers*; $\mathbf{C}$ the set of *complex numbers*. The nD Euclidean space will be denoted by $\mathbf{R}^n, n \in \mathbf{Z}$. We will use bold-faced letters, $\mathbf{x}, \mathbf{y}, \mathbf{k}$, for instance, to denote vectors or points in $\mathbf{R}^n$. The zero vector is denoted by $\mathbf{0}$. We do not distinguish the point $\mathbf{x}$ with the vector starting from the origin and ending at the point $\mathbf{x}$. The norm of a vector $\mathbf{x}$ is denoted by $|\mathbf{x}|$, standing for the quantity $(x_1^2 + \cdots + x_n^2)^{1/2}$, that is induced from the inner product $< \mathbf{x}, \mathbf{y} >= x_1y_1 + \cdots + x_ny_n$. We will use, in the standard sense, $L^p(\mathbf{R}^n)$ spaces, $1 \le p \le \infty$, and $L^1_{loc}(\mathbf{R}^n)$. By $B(\mathbf{x}_0, r)$ we denote the *open ball* in $\mathbf{R}^n$ centered at $\mathbf{x}_0$ with radius r. For any set A we denote by $|A|$ its *Lebesgue measure*. In this notation, $|B(\mathbf{x}_0, r)|$ stands for the volume of the ball $B(\mathbf{x}_0, r)$. The notation χ_E denotes the *characteristic function* of E that takes value 1 if $\mathbf{x} \in E$; and 0 if $\mathbf{x} \notin E$.

The concept *Lebesgue points* is needed in order to understand the theorems in §2. A point $\mathbf{x}$ is said to be a *Lebesgue point* of $f \in L^1_{loc}$, if

$$\lim_{r\to 0} \frac{1}{|B(\mathbf{x}, r)|} \int_{B(\mathbf{x},r)} |f(\mathbf{y}) - f(\mathbf{x})| d\mathbf{y} = 0.$$

It is easy to see that all continuous points of a function in L^1_{loc} are Lebesgue points. It is a striking fact that for any function in L^1_{loc}, and therefore any in $L^p, 1 \le p \le \infty$, almost all points are Lebesgue points of the function (see [12]).

For $f \in L^1$, the *Fourier transform* of f, denoted by $\mathcal{F}(f)$ or $\hat{f}$, is defined by

$$\mathcal{F}(f)(\xi) = \hat{f}(\xi) = \int_{\mathbf{R}^n} e^{-i<\mathbf{x},\xi>} f(\mathbf{x}) d\mathbf{x}.$$

For $f \in L^1 \cap L^2$, there holds the Plancherel relation:

$$\|f\|_2 = (2\pi)^{-(n/2)} \|\hat{f}\|_2.$$

This enables us to extend Fourier transform, from its definition on the dense subspace $L^1 \cap L^2$ of L^2, to the whole space L^2, in which the Plancherel relation remains to hold. The Plancherel relation implies that the Fourier transform in L^2 is invertible. The inverse is defined to be the inverse Fourier transform denoted by $\mathcal{F}^{-1}$.

The following simple properties of Fourier transformation will be used:

For $f \in L^1$, we have

(i) $\hat{f}$ is bounded and absolutely continuous in the whole space $\mathbf{R}^n$.

(ii) $\lim_{\xi\to\infty} \hat{f}(\xi) = 0$ (known as "Riemann-Lebesgue Theorem").

(iii)

$$\int_{\mathbf{R}^n} f(\mathbf{x})d\mathbf{x} = \hat{f}(\mathbf{0}). \tag{1}$$

(iv)

$$\mathcal{F}(e^{i<\mathbf{k},\cdot>}f(\cdot))(\xi) = \hat{f}(\xi - \mathbf{k}). \tag{2}$$

(v) Rotation on $\mathbf{R}^n$ is commutative with Fourier transformation, i.e.

$$\mathcal{F}(f(\rho(\cdot)))(\xi) = \hat{f}(\rho\xi), \tag{3}$$

where $\rho\xi$ is the image of ξ under the rotation ρ. If we use the notation $f^\rho(\mathbf{x}) = f(\rho\mathbf{x})$, the above relation may be written

$$\mathcal{F}(f^\rho) = (\mathcal{F}f)^\rho.$$

(vi) $\mathcal{F}(e^{-(1/2)|\cdot|^2})(\xi) = (2\pi)^{n/2}e^{-(1/2)|\xi|^2}$.

(vii) Fourier transform of a radial function (see the definition below) is radial ([12]).

A function $\phi \in L^2$ is said to be a *wavelet function*, or a *wavelet*, if

$$C_\phi = \frac{1}{\omega_{n-1}}\int_{\mathbf{R}^n} \frac{|\hat{\phi}(\xi)|^2}{|\xi|^n}d\xi < \infty, \tag{4}$$

where ω_{n-1} is the area of the $(n-1)$-dimensional unit sphere in $\mathbf{R}^n$. The relation (4) is called the *admissibility condition*. If ϕ also belongs to L^1, then the admissibility condition (4) implies

$$\int_{\mathbf{R}^n} \phi(\mathbf{x})d\mathbf{x} = 0. \tag{5}$$

A function satisfying (5) is said to have *vanishing moment* or *zero momentum*. The Fourier transformation relation (3) implies that the vanishing moment condition is equivalent to $\hat{\phi}(0) = 0$. Note that zero momentum is a necessary condition of the admissibility condition of a wavelet. The latter is a stronger condition that plays a crucial role for reconstructing the signal from its wavelet transform ([4,8]).

For a wavelet function ϕ we introduce the dilated translation

$$\begin{aligned} \phi_{a,\mathbf{b}}(\mathbf{x}) &= a^{-n}\phi((1/a)(\mathbf{x}-\mathbf{b})), \\ \phi_a(\mathbf{x}) &= \phi_{a,\mathbf{0}}(\mathbf{x}) = a^{-n}\phi((1/a)(\mathbf{x})),\ a > 0, \end{aligned} \tag{6}$$

where a^{-n} is the L^1-normalization factor. Indeed,

$$\int_{\mathbf{R}^n} \phi_{a,\mathbf{b}}(\mathbf{x})d\mathbf{x} = \int_{\mathbf{R}^n} \phi(\mathbf{x})d\mathbf{x}.$$

The *continuous wavelet transform* (CWT) of $f \in L^2$ associated with ϕ is defined by

$$(W_\phi f)(a,\mathbf{b}) =< \phi_{a,\mathbf{b}}, f >= \frac{1}{a^n}\int_{\mathbf{R}^n} \phi(1/a(\mathbf{x}-\mathbf{b}))\overline{f(\mathbf{x})}d\mathbf{x}, \tag{7}$$

where $\overline{f}$ denotes the complex conjugate of f. Recall that the convolution between two functions f and g is defined to be

$$f * g(\mathbf{b}) = \int_{\mathbf{R}^n} f(\mathbf{b}-\mathbf{x})g(\mathbf{x})d\mathbf{x} = \int_{\mathbf{R}^n} f(\mathbf{x})g(\mathbf{b}-\mathbf{x})d\mathbf{x}.$$

In this notation, the CWT may be rewritten

$$(W_\phi f)(a,\mathbf{b}) = (\phi_{a,\mathbf{0}} * \tilde{f})(\mathbf{b}), \tag{8}$$

where we used $\tilde{f}(\mathbf{x}) = \overline{f}(-\mathbf{x})$, which is called the *reflection* of f. In below we will only concern real-valued signal functions f and hence $\tilde{f}(\mathbf{x}) = f(-\mathbf{x})$.

The inversion formula reads

$$f(\mathbf{x}) = < f, \delta_{\mathbf{x}} > = \frac{1}{C_\phi} \int_0^\infty \int_{\mathbf{R}^n} W_\phi f(a, \mathbf{b}) \phi_{a,\mathbf{b}}(\mathbf{x}) \frac{da}{a} d\mathbf{b}.$$

In comparison with the Fourier series terminology, the CWT values $W_\phi f$ may be called the *coefficients* of f. The inversion formula amounts to reconstruct the original signal function f from its coefficients. Crucial to this study are those incorporated with a rotation effect in the transformation. Set

$$\phi_{a,\rho,\mathbf{b}}(\mathbf{x}) = a^{-n} \phi(a^{-1} \rho(\mathbf{x} - \mathbf{b})), \tag{9}$$

where ρ is a rotation of $\mathbf{R}^n$. The dilation is L^1-normalized:

$$\int_{\mathbf{R}^n} \phi_{a,\rho,\mathbf{b}}(\mathbf{x}) d\mathbf{x} = \int_{\mathbf{R}^n} \phi(\mathbf{x}) d\mathbf{x}.$$

We accordingly form the associated CWT

$$\begin{aligned} (W_\phi f)(a, \rho, \mathbf{b}) &= < f, \phi_{a,\rho,\mathbf{b}} > \\ &= a^{-n} \int_{\mathbf{R}^n} f(\mathbf{x}) \phi(a^{-1} \rho(\mathbf{x} - \mathbf{b})) d\mathbf{x}. \end{aligned} \tag{10}$$

In the convolution notation, it is

$$(W_\phi f)(a, \rho, \mathbf{b}) = f * \tilde{\phi}_{a,\rho,0}(\mathbf{b}). \tag{11}$$

In this case the inversion formula is

$$f(\mathbf{x}) = \frac{1}{C_\phi} \int_0^\infty \int_{SO(n)} \int_{\mathbf{R}^n} W_\phi f(a, \rho, \mathbf{b}) \phi_{a,\rho,\mathbf{b}}(\mathbf{x}) \frac{da}{a} d\mathbf{b} d\rho. \tag{12}$$

In applications we deal with the $2D$ case where the integration $\int_{S0(n)} \cdots d\rho$ reduces to $1/(2\pi) \int_0^{2\pi} \cdots d\theta$. Some commonly used wavelet functions are radial functions. In contrast, however, the $2D$ isotropy Morlet wavelet, or simply $2D$ Morlet wavelet is given by

$$\phi(\mathbf{x}) = e^{i<\mathbf{k},\mathbf{x}>} e^{-(1/2)|\mathbf{x}|^2} - e^{-(1/2)|\mathbf{k}|^2} e^{-(1/2)|\mathbf{x}|^2}, \tag{13}$$

where $\mathbf{k}$ is a fixed vector, is not radial, but sensitive to rotation. The rotation only has effect on the factor $e^{i<\mathbf{k},\mathbf{x}>}$.

It is easy to show that the Fourier transform of the Morlet wavelet is

$$\hat{\phi}(\xi) = 2\pi (e^{-(1/2)|\xi - \mathbf{k}|^2} - e^{-(1/2)|\mathbf{k}|^2} e^{-(1/2)|\xi|^2}). \tag{14}$$

We sometimes call the $2D$ Morlet wavelet *rotational* Morlet wavelet to emphasis the fact that it is incorporated with a "rotationally"-sensitive factor $e^{i<\mathbf{k},\mathbf{x}>}$. We also use the terminology *rotated* wavelet that means that instead of $\phi(\mathbf{x})$, it is the function $\phi^\rho(\mathbf{x}) = \phi(\rho \mathbf{x})$ that is concerned. For example, the rotational Morlet wavelet is the Morlet wavelet itself, but the rotated Morlet wavelet by ρ stands for

$$\begin{aligned} \phi^\rho(\mathbf{x}) &= e^{i<\mathbf{k},\rho\mathbf{x}>} e^{-(1/2)|\rho\mathbf{x}|^2} - e^{-(1/2)|\mathbf{k}|^2} e^{-(1/2)|\rho\mathbf{x}|^2} \\ &= e^{i<\mathbf{k},\rho\mathbf{x}>} e^{-(1/2)|\mathbf{x}|^2} - e^{-(1/2)|\mathbf{k}|^2} e^{-(1/2)|\mathbf{x}|^2}. \end{aligned} \tag{15}$$

It is often convenient to shift the rotation from $\mathbf{x}$ to $\mathbf{k}$, as

$$e^{i<\mathbf{k},\rho\mathbf{x}>} = e^{i<\rho^{-1}\mathbf{k},\rho^{-1}\rho\mathbf{x}>} = e^{i<\rho^{-1}\mathbf{k},\mathbf{x}>}.$$

The following section is devoted to establishment of a mathematical model for the rotational Morlet wavelet in extracting edge singularities.

To end this section we note that there are different ways to normalize the dilation of ϕ. For instance,

$$\phi_a(\mathbf{x}) = a^{-n} \phi(\mathbf{x}/a), \ \phi_a^{(2)}(\mathbf{x}) = a^{-n/2} \phi(\mathbf{x}/a).$$

The first (without a superscript, the same for the associated CWT) preserves the L^1-norm, while the second preserves the L^2-norm of the dilated ϕ.

By incorporating the rotation effect, in the L^2-normalization case, for instance, we have

$$\phi_{a,\rho,\mathbf{b}}^{(2)}(\mathbf{x}) = a^{-n/2} \phi(a^{-1} \rho(\mathbf{x} - \mathbf{b})),$$

$$(W_\phi^{(2)} f)(a,\rho,\mathbf{b}) =< f, \phi^{(2)}_{a,\rho,\mathbf{b}} >$$
$$= a^{-n/2}\int_{\mathbf{R}^n} f(\mathbf{x})\phi(a^{-1}\rho(\mathbf{x}-\mathbf{b}))d\mathbf{x} = \tilde{f} * \phi^{(2)}_{a,\rho,0}(\mathbf{b}).$$

The inversion formula correspondingly becomes

$$f(\mathbf{x}) = \frac{1}{C_\phi}\int_0^\infty \int_{SO(n)}\int_{\mathbf{R}^n} W_\phi^{(2)} f(a,\rho,\mathbf{b})\phi^{(2)}_{a,\rho,\mathbf{b}}(\mathbf{x})\frac{da}{a^{n+1}}d\mathbf{b}d\rho.$$

There is, in fact, no essential difference between different normalizations, as we have, for instance,

$$W_\phi f(a,\rho,\mathbf{b}) = \frac{1}{a^{n/2}} W_\phi^{(2)} f(a,\rho,\mathbf{b}).$$

3. **A Discussion On Mathematical Formulation of Edge Extraction.** This section is restricted to the $2D$ case and deals with only the $2D$ Morlet wavelet. Wavelet analysis involves two basic steps: One is wavelet transformation (WT), and the other is "adding up" the WTs to recover the original signal, known as inversion formula. We are to explore a method for edge detection in which only the CWT part will be used while the inversion formula for reconstruction of the original image signal will not be concerned.

To solve a practical problem one first formulates mathematical concepts and thus a theory as well based on anesthetics and metaphysics. The validity of the theory will be justified by its applicability. To formulate a mathematical theory for edge detection we first encounter the question: What is *edge*? The very naive answer would be: Edge is a piece of line segment or a piece of curve, and, in terms of a signal function, the function values on the piece of line segment or curve are significantly different from those elsewhere. Such a descriptive definition of edge immediately runs into difficulty due to the following two observations. First, lines and curves as edge are Lebesgue null sets. Lebesgue integral values, on the other hand, are not sensitive with function values on null sets: Even if we change the function values in a null set to be all ∞, the evaluation of the integral will still remain the same. So, edge information cannot be reflected by the CWT values. The second observation is based on Theorem 3.1 and 3.2 in below. They say that the CWT values at almost all points, not necessarily away from the edge points, will exhibit the same small magnitude if the parameter a is close to zero. So the edge will not be detected.

In below we split the $2D$ Morlet wavelet (13) into

$$\phi(\mathbf{x}) = \phi_M(\mathbf{x}) + \phi_E(\mathbf{x}),$$

where

$$\begin{aligned} \phi_M(\mathbf{x}) &= e^{i<\mathbf{k},\mathbf{x}>}e^{-(1/2)|\mathbf{x}|^2} \\ \phi_E(\mathbf{x}) &= -e^{-(1/2)|\mathbf{k}|^2}e^{-(1/2)|\mathbf{x}|^2}. \end{aligned} \tag{16}$$

Their Fourier transforms are

$$\begin{aligned} \hat{\phi_M}(\xi) &= 2\pi e^{-(1/2)|\xi-\mathbf{k}|^2} \\ \hat{\phi_E}(\xi) &= -2\pi e^{-(1/2)|\mathbf{k}|^2}e^{-(1/2)|\xi|^2}, \end{aligned} \tag{17}$$

respectively. We shall call ϕ_M the rotational part, and ϕ_E the radial part. In (10), with ϕ_M and ϕ_E in place of ϕ, we will still call the obtained, with a little abuse of the terminology CWT, the CWT associated with ϕ_M and ϕ_E, respectively.

Theorem 3.1. *Let $f \in L^p(\mathbf{R}^2), 1 \le p \le \infty$. Then for any rotation $\rho \in SO(2)$ and almost all $\mathbf{b} \in \mathbf{R}^2$, we have*

$$\lim_{a\to 0}(W_\phi f)(a,\rho,\mathbf{b}) = 0.$$

Theorem 3.2. *Let $f \in L^p(\mathbf{R}^2), 1 \le p \le \infty$. Then for every rotation $\rho \in SO(2)$ and almost all $\mathbf{b} \in \mathbf{R}^2$, we have*

$$\begin{aligned} \lim_{a\to 0}(W_{\phi_M} f)(a,\rho,\mathbf{b}) &= \lim_{a\to 0}(W^{(1)}_{\phi_E} f)(a,\rho,\mathbf{b}) \\ &= 2\pi e^{-(1/2)|\mathbf{k}|^2} f(\mathbf{b}). \end{aligned} \tag{18}$$

Theorem 3.1 and 3.2 are immediate consequences of the general result: If there exists a radial function Ψ in L^1 dominating $\Phi : |\Phi(\mathbf{x})| \le \Psi(\mathbf{x})$, then, for any function $f \in L^p, 1 \le p \le \infty$, we have

$$\lim_{a\to 0}\Phi_a * f(\mathbf{b}) = \hat{\Phi}(\mathbf{0})f(\mathbf{b})$$

at least at all Lebesgue points $\mathbf{b}$ of f (Theorem 1.25, Chapter 1, [12]). Since almost all points of $\mathbf{R}^n$ are Lebesgue points of f, the above holds for almost all $\mathbf{b}$.

We will adopt the notation $\phi(a,\rho,\mathbf{x}) = (\phi^\rho)_a(\mathbf{x})$, where $\phi^\rho(\mathbf{x}) = \phi(\rho\mathbf{x})$ is the rotated wavelet, and $W_\phi f(a,\rho,\mathbf{b}) = (\tilde{f} * \phi_a^\rho)(\mathbf{b})$. When $\hat{\Phi}(\mathbf{0}) = 1$, then the sequence $(\Phi)_a, a \to 0$, is called *an approximation to the identity* whose role is the same as the Dirac δ-distribution. In below we will refer this result several times. For convenience, no matter whether it is with the condition $\hat{\Phi}(\mathbf{0}) = 1$, we will call it an *approximation net*. The theorems are proved by invoking the above mentioned general result in [12] to $(\phi^\rho)\tilde{}, (\phi_M^\rho)\tilde{}$ and $(\phi_E^\rho)\tilde{}$. Note that the functions $(\phi^\rho), (\phi_M^\rho)$ and (ϕ_E^ρ) are identical with their reflections.

Therefore, on Lebesgue points and for large $\mathbf{k}$ all CWT values exhibit small magnitudes in the procedure $a \to 0$ and neither ϕ, nor ϕ_M, nor ϕ_E are useful to detect edge singularities represented by null sets in the process $a \to 0$.

These observations are not disappointing. They suggest that to give a mathematical formulation for edge detection one would proceed with one of the following two strategies.

One is to give up the assumption that the edge information is carried by a null set, but by a region. In a photo at the places where the grey scale level changes rapidly is considered to be the edge. Theorem 3.1 and 3.2 also tell that in edge detection, when using CWT the parameter, a should not be chosen too small.

Alternatively, one may choose to stick on the intuition that "edge" is a piece of line or curve segment of Lebesgue measure zero. This very naive starting point will be regarded as the "ideal situation" in the sequel. The rest of the paper will be devoted to a theory for this case based on a particular application of the Dirac δ-distribution belonging to the category of Radon measures. The formulation will clearly indicate why rotational wavelets work well in edge detection and how to imply it in relation to selection of the parameters.

4. **The Radon Measure Formulation: The Ideal Situation.** In order to pursue a study of edge extraction under the assumption that the point set of edge is contained in a Lebesgue null set, the concept of Radon measure would be unavoidable. The point is that on Lebesgue null sets integrals (representing Radon measure as functionals) are not necessarily zero. In the language of measure theory, this amounts to say that measures we use should not be absolutely continuous with respect to the Lebesgue measure.

We briefly introduce Radon measures in $\mathbf{R}^n$. A *Radon measure* is a Borel measure with finite values on compact sets. By C_c we denote the set of complex-value continuous functions with compact support, and by C_c^+ the functions in C_c that only take non-negative values. Functions in C_c are called test functions.

Definition 4.1. *A complex-valued function I defined on C_c is said to be a non-negative linear functional, or simply functional, if for all $f, g \in C_c$ and $\alpha \in \mathbf{C}$*

- $I(f+g) = I(f) + I(g)$ *(additive);*
- $I(\alpha f) = \alpha I(f)$ *(homogeneous); and*
- $I(f) \geq 0$, *if $f \in C_c^+$. (non-negative).*

It can be shown that every non-negative functional I on C_c induces a Radon measure such that

$$I(f) = \int_{\mathbf{R}^n} f d\mu_I, \qquad f \in C_c.$$

On the other hand, every Radon measure induces a non-negative linear functional on C_c. The functional introduced by a Radon measure $d\mu$ is given by

$$< f, d\mu >= \int_{\mathbf{R}^n} f(\mathbf{x}) d\mu(\mathbf{x}).$$

We say that a functional is *supported in a closed set B*, if it vanishes on all test functions whose compact support has an empty intersection with B. In the theory of generalized functions a Radon measure is a tempered distribution on the class of infinitely differentiable functions with compact support.

What is in mind with use of Radon measures is functionals with mass evenly distributed on closed curves or surfaces. Let δ be the standard Dirac distribution with the characteristic property

$$\int_{\mathbf{R}^n} f(\mathbf{y})\delta(\mathbf{x}-\mathbf{y})d\mathbf{y} = f(\mathbf{x}), \qquad f \in C_c.$$

Now let $G(\mathbf{x}) = 0$ represent a curve in $\mathbf{R}^2$ or a surface in $\mathbf{R}^n$. We claim, without giving a proof, that the measure $d\mu(\mathbf{x}) = \delta(G(\mathbf{x}))d\mathbf{x}$ is the Radon measure for the functional supported and evenly distributed on the curve or surface $G = G(\mathbf{x}) = 0$:

$$< f, d\mu >= \int_G f(\mathbf{x}) ds(\mathbf{x}),$$

where $ds(\mathbf{x})$ denotes the arc length or surface area measure on G. Distributions of this kind are used in obtaining a fundamental solution of the standard wave equation in $\mathbf{R}^n$ in which the surfaces are spheres corresponding to $G(\mathbf{x}) = |\mathbf{x} - \mathbf{a}| - r = 0$.

The rest of the section will be restricted to $\mathbf{R}^2$, and the measure $\delta(G(\mathbf{x}))d\mathbf{x}$ represents the signal function. In our model G is the edge to be detected on which the signal (the grey scale level) is ∞, and, away from the edge, zero. We will be working on the Radon measure $d\mu(x_1, x_2) = \delta((1/\sqrt{2})(x_1 - x_2))dx_1dx_2$ as a sample that represents the straight line $L = \{(x_1, x_2) \mid x_1 = x_2\}$. In the spirit of the following computation the proofs in Section 7 are proceeded. For any f in C_c, by performing the rotation $x_1 = (1/\sqrt{2})(y_1 + y_2), x_2 = (1/\sqrt{2})(y_1 - y_2)$, followed by iterated integration, we have

$$\begin{aligned} < f, d\mu >= & \int_{\mathbf{R}\times\mathbf{R}} f(x_1, x_2)\delta(\frac{1}{\sqrt{2}}(x_1 - x_2))dx_1dx_2 \\ = & \int_{\mathbf{R}}(\int_{\mathbf{R}} f(\frac{1}{\sqrt{2}}(y_1 + y_2), \frac{1}{\sqrt{2}}(y_1 - y_2))\delta(y_2)dy_2dy_1 \\ = & \int_{\mathbf{R}} f(\frac{1}{\sqrt{2}}y_1, \frac{1}{\sqrt{2}}y_1)dy_1 \\ = & \sqrt{2}\int_{\mathbf{R}} f(y_1, y_1)dy_1 \\ = & \int_l f(x_1, x_2)ds, \end{aligned}$$

where ds is the arc-length measure on L. This shows that the functional reduces to the line integral along the line L. As consequence, if the support of f is away from the line $\{x_1 = x_2\}$, then the value of the functional on f is identical with zero. Thus the functional is supported on the line. We can proceed with a general line $\{ax_1 + bx_1 + c = 0\}$ and obtain the same conclusion. With these preparations we now state our main theorem that, for simplicity, is for the particular line L. The conclusion, however, can be adapted to general cases.

Theorem 4.1. *Denote* $l(\mathbf{x}) = \delta((1/\sqrt{2})(x_1 - x_2))$. *Then*

(i) *In the functional sense,*

$$\begin{aligned} \lim_{a\to 0}(W_{\phi_M} l)(a, \rho, \mathbf{b}) &= \lim_{a\to 0}(W_{\phi_E} l)(a, \rho, \mathbf{b}) \\ &= 2\pi e^{-(1/2)|\mathbf{k}|^2} l(\mathbf{b}) \end{aligned}$$

(ii) *For any* $a > 0$ *and* $\mathbf{b} \in \mathbf{R}^2$, *in the pointwise sense,*

$$(W^{(2)}_{\phi_E} l)(a, \rho, \mathbf{b}) = \sqrt{2\pi} e^{-(1/2)|\mathbf{k}|^2} e^{-d^2/(2a^2)},$$

where d *is the distance of* $\mathbf{b}$ *to the line* $L = \{x_1 = x_2\}$.

(iii) *For any* $a > 0$ *and* $\mathbf{b} \in \mathbf{R}^2$, *in the pointwise sense,*

$$\begin{aligned} &(W^{(2)}_{\phi_M} l)(a, \rho, \mathbf{b}) = \\ &= \sqrt{2\pi} e^{i(\cos\theta)\mathbf{k}d/a} e^{-d^2/(2a^2)} e^{-(1/2)(\sin\theta|\mathbf{k}|)^2}, \end{aligned}$$

where as in (ii), d *is the distance from* $\mathbf{b}$ *to* L, *and* θ *is the angle between the argument* $\rho^{-1}\mathbf{k}$ *and that of the normal of the line:* $\theta = \arg\rho^{-1}\mathbf{k} - 3\pi/4$.

The proof is postponed to §7 where we prove the nD analogy of the theorem for a class of wavelets generalizing the Morlet wavelet. We now give an explanation of the theorem.

Owing to the characteristic property of l, for the points $\mathbf{b}$ not on the line L the assertion (i) and those in (ii) and (iii) corresponding to $d > 0$ show that neither the CWTs of $W^{(2)}_{\phi_M} l$, nor that of $W^{(2)}_{\phi_E} l$ exhibit significant values in the procedure $a \to 0$. In fact, whenever $d > 0$ the values tend to zero for any fixed $\mathbf{k}$ and ρ. If a point $\mathbf{b}$ is on the line, corresponding to $d = 0$ in (ii) and (iii), then the L^2-normalized CWT values are not sensitive to the parameter a. As a matter of fact, the norm of the CWTs are independent of a. The CWT values for ϕ_E are independent of the rotation parameter, while in contrast, the CWT values for ϕ_M significantly depend on the rotation: when $\rho^{-1}\mathbf{k}$ is orthogonal to the line L, that is $\theta = 0$, the norms of the corresponding CWT values take the maximal value $\sqrt{2\pi}$, and in general, the norm of the CWT value for ϕ_M is the greater the closer from $\arg \rho^{-1}\mathbf{k}$ to $3\pi/4$.

When the edge is a general line $\{ax_1 + bx_1 + c = 0\}$, we can reduce it to the case considered in the theorem and obtain the same conclusions. Now assume that the edge consists of several lines each corresponding to a Radon measure $\delta(a_k x_1 + b_k x_2)d\mathbf{x}, k = 1, 2, ..., m$. Then the signal function is the summation of the Radon measures. From the above analysis we see that when the direction of $\rho^{-1}\mathbf{k}$ is orthogonal, or nearly orthogonal with one of the lines, then CWT values on that particular line will exhibit greater magnitudes, and on the other lines and other places away from the lines the CWT values are comparatively smaller. This suggests how one should do edge detection by making use of the rotational Morlet wavelet ϕ_M. Suppose now we want to detect a face boundary (edge) in a photo that is not a straight line but a closed oval-shaped curve. For an arbitrary chosen rotation ρ, as above observed, in the procedure $a \to 0$ we extract out from the maximal magnitudes of the CWTs $(W^{(2)}_{\phi_M} l)(a, \rho, \mathbf{b})$ the boundary portions that are orthogonal or nearly orthogonal with the rotated vector $\rho^{-1}\mathbf{k}$. Then by altering the rotation ρ, for instance choosing six or eight evenly distributed directions of the whole cycle, we obtain boundary portions in almost all directions. By connecting them we obtain the whole face boundary.

Suppose now we deal with a photo consisting of locally only two grey scale levels and the edge is the dividing curve of the two regions of different grey scale levels, then the above method should be modified. In this model one uses $h(\mathbf{x}) = H(< \mathbf{a}, \mathbf{x} >)$ in place of $l(\mathbf{x}) = \delta(< \mathbf{a}, \mathbf{x} >)$ in the CWT formula, where H is the Heaviside function: $H(x) = 1, x > 0$; and $H(x) = 0, x \le 0$, and $\mathbf{a}$ the unit normal of the edge. Then we take the directional derivative to the CWT of h along the unit vector $\mathbf{p} = (p_1, p_2)$ orthogonal to $\rho^{-1}\mathbf{k}$. The derivative will be passed on to $H(< \mathbf{a}, \mathbf{x} >)$. Invoking the result $dH/dx = \delta$ in the generalized function theory, we have

$$(\partial/\partial\mathbf{p})H(< \mathbf{a}, \mathbf{x} >) =< \mathbf{a}, \mathbf{p} > \delta(< \mathbf{a}, \mathbf{x} >).$$

This introduces the relation

$$(\partial/\partial\mathbf{p})[(W^{(2)}_{\phi_M} h)(a, \rho, \mathbf{b})] =< \mathbf{a}, \mathbf{p} > (W^{(2)}_{\phi_M} l)(a, \rho, \mathbf{b}).$$

The norm of the directional derivative takes the maximal value when $\mathbf{a} = \pm\mathbf{p}$, that is when $\mathbf{a}$ is orthogonal with $\rho^{-1}\mathbf{k}$, consistent with the conclusion of Theorem 4.1. In practice, when detecting edges in a photo with locally two grey scale levels, for the selected $\mathbf{k}$, and then ρ, by finding the maximal magnitudes of the norm of the directional derivatives along $\mathbf{p}$ of the CWT of the signal function we extract the edge pieces orthogonal with the direction $\rho^{-1}\mathbf{k}$. Detailed analysis will show that the above is an analogy to Canny's gradient method but combined with the rotation factor. In practice we use other variations according to different natures of problems. Note that we are now dealing with the ideal situation and the condition $a \to 0$ corresponds to the assumption that the width of the edge is zero. In the real life situation the parameter a should not tend to zero but be restricted to a certain band comparable to the width of the edge. The real life situation and the variations of the method will be discussed in a separate paper. Applications of rotational wavelets in edge detection may also be found in [11-13].

5. **Examples of Implementation of the Theory.** When applying the method established through Theorem 4.1, discretization of the integral defining the CWTs will be involved.

5.1. **Implementation of wavelet transform on images.** Application of rotational wavelets includes three steps:

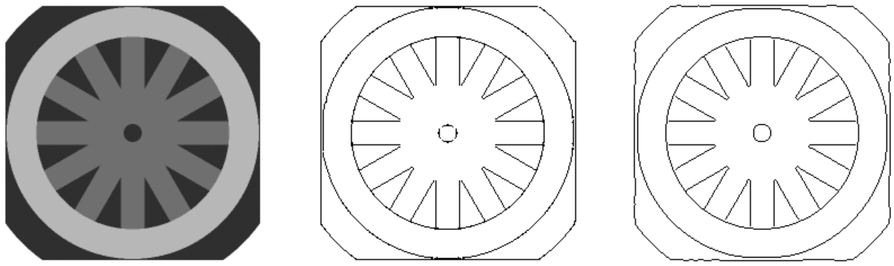
5.1.1. *Discretization of the integral CWTs.* Assume $f(x, y)$ is an image of size M_1 in width and M_2 in height. Direct discretization of CWT of $f(x, y)$ is like this:

$$W^{(2)}_{\phi}(f)(a, \rho_i, b) = a^{-1} \sum_{u_1=b_2-N/2}^{b_2+N/2} \sum_{u_2=b_1-N/2}^{b_1+N/2} G(u_1, u_2)$$

where

$$G(u_1, u_2) := f(u_1, u_2)\phi_M((a^{-1}\rho_i(u_1 - b_1, u_2 - b_2)),$$

N is the range of convolution, ρ_i is rotation of degrees $\theta_i = 0, \pi/4, \pi/2, 3\pi/4$.The parameters $N, a, (k_1 k_2)$ are chosen as below: the value of N is as small as possible and insures the restoration

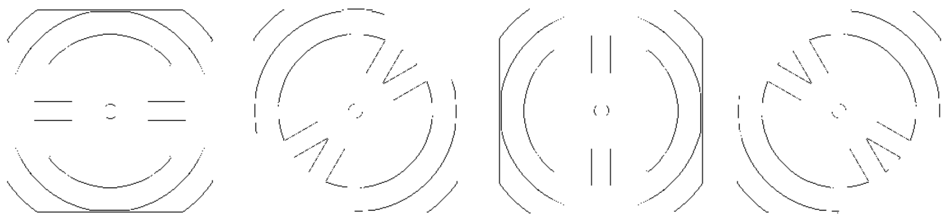


(a) Original wheel image

(b) Edges extracted by Morlet wavelet

(c) Edges extracted by Canny operator

FIGURE 1. Original image and edges extracted by rotational Morlet wavelet and Canny operator



(a) Edges of 0 degree

(b) Edges of 45 degree

(c) Edges of 90 degree

(d) Edges of 135 degree

FIGURE 2. Edges of four directions

of wavelet functions. By Theorem 4.1, when the width of the edge is zero, only when $a \to 0$ can edges be efficiently extracted. In practice, the width of the edges are not zero, and the values of a are chosen by experience or through experiments. Here we adopt the parameters $a = 0.125, (k_1, k_2) = (6.9, 0)$.

5.1.2. *Compute directional derivative of wavelet transform.* In practice, Images usually contain step edges, so by the explanation of Theorem 4.1, directional derivative should be applied on images first.

5.1.3. *Through threshold transformation to get edges.*

5.2. **Results of experiments.** We have tested three kinds of images: regular pattern image, portrait, and remote sensing image. They cover typical examples we usually come across. Because the Canny operator was considered to be the optimal, below we list some results derived from both the rotational Morlet wavelet and form the Canny operator for comparison. Fig.2 presents edges of four directions extracted from typical pattern image by rotational wavelet, which shows that rotational wavelet can efficiently extract edges of different direction and rotational wavelet is sensitive to direction. The comparisons in Fig.4 and Fig. 3 show that the edges extracted by the Morlet wavelet is not so fine in connectivity but more precise in location.

(a) Original Lenna image

(b) Edges extracted by Morlet wavelet

(c) Edges extracted by Canny operator

FIGURE 3. Original image and edges extracted by rotational Morlet wavelet and Canny operator

(a) Original remote sensing image

(b) Edges extracted by Morlet wavelet

(c) Edges extracted by Canny operator

FIGURE 4. Original image and edges extracted by rotational Morlet wavelet and Canny operator

6. **Conclusion.** Rotational Morlet wavelet can efficiently extract edges of different directions. It therefore is suitable for the cases such as remote sensing images and pattern images. The edges are extracted when the parameter a is chosen to be comparable with width of the edges. The resolution is precise in location.

7. **Generalizations of the Morlet Wavelet.** The task of this section is to prove the theorems of Section 3 and 4 in a more general context. We first generalize the Morlet wavelet. Unless otherwise stated, throughout this section Φ will be an arbitrary but fixed function, satisfying the conditions:

(i)

$$\Phi \in L^1 \cap L^2.$$

(ii) There exists a radial function Ψ such that

$$|\Phi(\mathbf{x})| \leq \Psi(\mathbf{x}).$$

(iii)

$$\int_{\mathbf{R}^n} (1 + |\mathbf{x}|)\Psi(\mathbf{x}) d\mathbf{x} < \infty.$$

(iv)

$$\int_{\mathbf{R}^n} \Phi(\mathbf{x}) d\mathbf{x} = \hat{\Phi}(\mathbf{0}) \neq 0.$$

We have the following

Theorem 7.1. *For any vector $\mathbf{k} \in \mathbf{R}^n$, the function*

$$\phi(\mathbf{x}) = \hat{\Phi}(\mathbf{0}) e^{i<\mathbf{k},\mathbf{x}>} \Phi(\mathbf{x}) - \hat{\Phi}(-\mathbf{k}) \Phi(\mathbf{x})$$

is a wavelet.

Proof: It is obvious that $\phi \in L^1 \cap L^2$. It suffices to verify that ϕ satisfies the admissibility condition (4). Based on the Fourier transformation relations cited in §1, we have

$$\hat{\phi}(\xi) = \hat{\Phi}(\mathbf{0}) \hat{\Phi}(\xi - \mathbf{k}) - \hat{\Phi}(-\mathbf{k}) \hat{\Phi}(\xi).$$

We first have

$$\begin{aligned} \int_{|\xi| \geq 1} \frac{|\hat{\phi}(\xi)|^2}{|\xi|^n} d\xi &\leq \int_{|\xi| \geq 1} |\hat{\phi}(\xi)|^2 d\xi \leq \int_{\mathbf{R}^n} |\hat{\phi}(\xi)|^2 d\xi \\ &= \|\hat{\phi}\|_2^2 = (2\pi)^{n/2} \|\phi\|_2^2 < \infty. \end{aligned}$$

To show the integrability at the origin we notice that

$$\begin{aligned} \hat{\phi}(\xi) &= \hat{\Phi}(\mathbf{0})(\hat{\Phi}(\xi - \mathbf{k}) - \hat{\Phi}(-\mathbf{k})) + \hat{\Phi}(-\mathbf{k})(\hat{\Phi}(\mathbf{0}) - \hat{\Phi}(\xi)) \\ &= \hat{\Phi}(\mathbf{0}) g_{-\mathbf{k}}(\xi) - \hat{\Phi}(-\mathbf{k}) g_{\mathbf{0}}(\xi), \end{aligned}$$

where for any $\mathbf{j} \in \mathbf{R}^n$, we use the notation

$$g_{\mathbf{j}}(\xi) = \hat{\Phi}(\xi + \mathbf{j}) - \hat{\Phi}(\mathbf{j}).$$

Owing to the elementary inequality $|a+b|^2 \leq 2|a|^2 + 2|b|^2$,we have

$$\begin{aligned} \int_{|\xi| \leq 1} \frac{|\hat{\phi}(\xi)|^2}{|\xi|^n} d\xi &\leq 2|\hat{\Phi}(\mathbf{0})|^2 \int_{|\xi| \leq 1} \frac{|g_{-\mathbf{k}}(\xi)|^2}{|\xi|^n} d\xi \\ &+ 2|\hat{\Phi}(-\mathbf{k})|^2 \int_{|\xi| \leq 1} \frac{|g_{\mathbf{0}}(\xi)|^2}{|\xi|^n} d\xi. \end{aligned}$$

We now show that for any $\mathbf{j}$,

$$|g_{\mathbf{j}}(\xi)| \leq C|\xi|,$$

where C is a constant depending only on Ψ. Indeed, owing to the condition (iii) and the Lebesgue dominated convergence theorem, $\hat{\Phi}$ has first order partial derivatives everywhere, and

$$\begin{aligned} |\nabla \hat{\Phi}(\xi)| &\leq \int_{\mathbf{R}^n} |\nabla e^{-i<\xi,\mathbf{x}>}| |\Phi(\mathbf{x})| d\mathbf{x} \\ &\leq \int_{\mathbf{R}^n} |\mathbf{x}| |\Psi(\mathbf{x})| d\mathbf{x} \\ &= C < \infty, \end{aligned}$$

where C depends on Ψ. Therefore,

$$|g_{\mathbf{j}}(\xi)| \leq \max\{|\nabla \hat{\Phi}|\} |\xi| \leq C|\xi|.$$

Now

$$\int_{|\xi| \leq 1} \frac{|g_{\mathbf{j}}(\xi)|^2}{|\xi|^n} d\xi \leq C \int_{|\xi| \leq 1} \frac{1}{|\xi|^{n-1}} d\xi < \infty$$

implies

$$\int_{|\xi| \leq 1} \frac{|\hat{\phi}(\xi)|^2}{|\xi|^n} d\xi < \infty.$$

The proof is complete.

We can construct rotational wavelets based on Theorem 7.1. For instance, the $2D$ and nD Morlet wavelets are constructed from the function $\Phi(\mathbf{x}) = e^{-(1/2)|\mathbf{x}|^2}$, the Gaussian density function. Indeed, in the case, $\Psi = \Phi$ and $\hat{\Phi}(\xi) = (2\pi)^{n/2} e^{-(1/2)|\xi|^2}$, and so $\hat{\Phi}(\mathbf{0}) = (2\pi)^{n/2}$. We therefore have

$$\phi(\mathbf{x}) = (2\pi)^{\frac{n}{2}} e^{i<\mathbf{k},\mathbf{x}>} e^{-\frac{1}{2}|\mathbf{x}|^2} - (2\pi)^{\frac{n}{2}} e^{-\frac{1}{2}|\mathbf{k}|^2} e^{-\frac{1}{2}|\mathbf{x}|^2}.$$

There is a great freedom to choose Φ under the conditions (i) to (iv), and one can establish a bank of rotational wavelets of the Morlet type. In particular, the commonly used kernel functions for approximation to the identity can all be used for Φ (See the examples to the end of this section.). Bear in mind, however, that Φ itself should not have zero momentum, for otherwise, the term incorporating the rotation factor will disappear. The results and the proofs in §3, and the theorem in §4 for the $2D$ case can all be extended to $\mathbf{R}^n$ without encountering essential difficulty. We have

Theorem 7.2. *Let $f \in L^p(\mathbf{R}^n), 1 \le p \le \infty$. Then for any rotation $\rho \in SO(n)$ and almost all $\mathbf{b} \in \mathbf{R}^n$, we have*

$$\lim_{a\to 0}(W_\phi f)(a,\rho,\mathbf{b}) = 0.$$

The proof invokes Theorem 1.25, Chapter 1, [12].

As before, we truncate ϕ into its rotational and radial parts:

$$\phi(\mathbf{x}) = \phi_M(\mathbf{x}) - \phi_E(\mathbf{x}),$$

where

$$\phi_M(\mathbf{x}) = \hat{\Phi}(\mathbf{0})e^{i<\mathbf{k},\mathbf{x}>}\Phi(\mathbf{x}) \quad \text{and} \quad \phi_E(\mathbf{x}) = \hat{\Phi}(-\mathbf{k})\Phi(\mathbf{x}).$$

Their Fourier transforms are

$$\hat{\phi_M}(\xi) = \hat{\Phi}(\mathbf{0})\hat{\Phi}(\xi - \mathbf{k}) \quad \text{and} \quad \hat{\phi_E}(\xi) = \hat{\Phi}(-\mathbf{k})\hat{\Phi}(\xi),$$

respectively. With the same proof as for Theorem 3.2, we have

Theorem 7.3. *Let $f \in L^p(\mathbf{R}^n), 1 \le p \le \infty$. Then for any rotation $\rho \in SO(n)$ and almost all $\mathbf{b} \in \mathbf{R}^n$, we have*

$$\begin{aligned}\lim_{a\to 0}(W_{\phi_M} f)(a,\rho,\mathbf{b}) &= \lim_{a\to 0}(W_{\phi_E} f)(a,\rho,\mathbf{b}) \\ &= \hat{\Phi}(\mathbf{0})\hat{\Phi}(-\mathbf{k})f(\mathbf{b}).\end{aligned}$$

By a $(n-1)$-dimensional hyperplane we mean the $(n-1)$-dimensional linear subspace in $\mathbf{R}^n$ satisfying the equation $< \mathbf{a},\mathbf{x} >= a_1x_1 + \cdots + a_nx_n = 0$, where $\mathbf{a} = (a_1, \cdots, a_n)$ is a unit normal of the hyperplane.

Theorem 7.4. *Let $\mathbf{a}$ be a unit vector. Denote by $l(\mathbf{x}) = \delta(< \mathbf{a},\mathbf{x} >)$ the Dirac distribution supported on the hyperplane $\{a_1x_1 + \cdots + a_nx_n = 0\}$. Assume that Φ is radial and satisfies the conditions* (i) *to* (iv). *Then we have*

(i) *In the functional sense,*

$$\begin{aligned}\lim_{a\to 0}(W_{\phi_M} l)(a,\rho,\mathbf{b}) &= \lim_{a\to 0}(W_{\phi_E} l)(a,\rho,\mathbf{b}) \\ &= \hat{\Phi}(\mathbf{0})\hat{\Phi}(\mathbf{k})l(\mathbf{b}).\end{aligned}$$

(ii) *For any $a > 0$ and $\mathbf{b} \in \mathbf{R}^n$, in the pointwise sense,*

$$\begin{aligned}a^{n/2-1}(W^{(2)}_{\phi_E} l)(a,\rho,\mathbf{b}) &= \hat{\Phi}(\mathbf{k})\mathcal{F}^{(n-1)}(\Phi_{(d/a)})(0) \\ &= \hat{\Phi}(\mathbf{k})\psi_{(d/a)}(0),\end{aligned}$$

where d is the distance from the point $\mathbf{b}$ to the hyperplane, the jth-component of a general vector $\mathbf{v}$ is denoted by $(\mathbf{v})_j$, where
$\mathbf{v}^{(n-1)} = ((\mathbf{v})_1, \cdots, (\mathbf{v})_{n-1}), \Phi_{(\mathbf{x})_n}(\mathbf{x}^{(n-1)}) = \Phi(\mathbf{x}), \mathcal{F}^{(n-1)}(\Phi_{(\mathbf{v})_n})(\xi^{(n-1)}) = \psi_{(\mathbf{v})_n}(|\xi^{(n-1)}|)$, and $\mathcal{F}^{(n-1)}$ stands for the Fourier transform in $\mathbf{R}^{(n-1)}$ (The existence of the function $\psi : [0,\infty) \to \mathbf{C}$ is owing to the property (vii) of Fourier transformation).

(iii) *For any $a > 0$ and $\mathbf{b} \in \mathbf{R}^n$, in the pointwise sense,*

$$a^{1-\frac{n}{2}}(W^{(2)}_{\phi_M} l)(a,\rho,\mathbf{b}) = \hat{\Phi}(\mathbf{0})e^{i\cos\theta|\mathbf{k}|\frac{d}{a}}\psi_{\frac{d}{a}}(\sin\theta|\mathbf{k}|),$$

where θ denotes the angle between the vector $\rho^{-1}\mathbf{k}$ and the normal of the hyperplane.

Proof: To prove (i) let f be any test function in C_c. Let ρ_0 be the rotation such that $\rho_0^{-1}\mathbf{a} = \mathbf{e}_n = (0, \cdots, 0, 1)$, and thus ρ_0 rotates the plane $x_n = 0$ to the plane $< \mathbf{a},\mathbf{x} >= 0$. Since Lebesgue measure is invariant under rotation, using the notation $f^{\rho_0}(\mathbf{x}) = f(\rho_0\mathbf{x})$, we have

$$
\begin{aligned}
& < (W_{\phi_M} l)(a, \rho, \cdot), f > = < (\phi_M^\rho)_a * \tilde{l}, f > \\
= & \int_{\mathbf{R}^n} (\phi_M^\rho)_a(\mathbf{x}) \int_{\mathbf{R}^n} f(\rho_0\mathbf{b})\delta(< \mathbf{a}, \mathbf{x} - \rho_0\mathbf{b} >) d\mathbf{b} d\mathbf{x} \\
= & \int_{\mathbf{R}^n} (\phi_M^\rho)_a(\mathbf{x}) \int_{\mathbf{R}^n} f^{\rho_0}(\mathbf{b})\delta(< \rho_0^{-1}\mathbf{a}, \rho_0^{-1}\mathbf{x} - \mathbf{b} >) d\mathbf{b} d\mathbf{x} \\
= & \int_{\mathbf{R}^n} (\phi_M^\rho)_a(\mathbf{x}) \int_{\mathbf{R}^n} f^{\rho_0}(\mathbf{b})\delta(< \mathbf{e}_n, \rho_0^{-1}\mathbf{x} - \mathbf{b} >) d\mathbf{b} d\mathbf{x} \\
= & \int_{\mathbf{R}^n} (\phi_M^\rho)_a(\mathbf{x}) \int_{\mathbf{R}^{n-1}} \Big(\int_{\mathbf{R}} f^{\rho_0}(b_1, \cdots, b_{n-1}, b_n) \\
& \delta((\rho_0^{-1}\mathbf{x})_n - b_n) db_n \Big) db_1 \cdots db_{n-1} d\mathbf{x} \\
= & \int_{\mathbf{R}^n} (\phi_M^\rho)_a(\mathbf{x}) \int_{\mathbf{R}^{n-1}} f^{\rho_0}(b_1, \cdots, b_{n-1}, (\rho_0^{-1}\mathbf{x})_n) db_1 \cdots db_{n-1} d\mathbf{x},
\end{aligned}
$$

where we denote $(\rho_0^{-1}\mathbf{x})_n$ the nth component of $\rho_0^{-1}\mathbf{x}$. Using the result of approximation net and the value of $\mathcal{F}(\phi_M)(0)$, the above is equal to

$$
\begin{aligned}
& \hat{\Phi}(\mathbf{0})\hat{\Phi}(-\mathbf{k}) \int_{\mathbf{R}^{n-1}} f^{\rho_0}(b_1, \cdots, b_{n-1}, 0) db_1 \cdots db_{n-1} \\
= & \hat{\Phi}(\mathbf{0})\hat{\Phi}(\mathbf{k}) < l, f > .
\end{aligned}
$$

The assertion for ϕ_E can be proved similarly.

Next we first prove (iii). Performing the translation $\mathbf{x}' = \mathbf{x} - \mathbf{b}$ in the integral and noting that $| < \mathbf{a}, \mathbf{b} > | = d$ is the distance from $\mathbf{b}$ to the hyperplane $< \mathbf{a}, \mathbf{x} >= 0, |\mathbf{a}| = 1$, we have

$$
\begin{aligned}
& < a^{1-n/2}(\phi_M^\rho)_a, l > \\
= & \hat{\Phi}(\mathbf{0}) a^{1-n} \int_{\mathbf{R}^n} e^{i < \rho^{-1}\mathbf{k}, \frac{\mathbf{x}-\mathbf{b}}{a} >} \Phi(\frac{\mathbf{x}-\mathbf{b}}{a}) \delta(< \mathbf{a}, \mathbf{x} >) d\mathbf{x} \\
= & \hat{\Phi}(\mathbf{0}) a^{1-n} \int_{\mathbf{R}^n} e^{i < \rho^{-1}\mathbf{k}, \frac{\mathbf{x}'}{a} >} \Phi(\frac{\mathbf{x}'}{a}) \delta(< \mathbf{a}, \mathbf{x}' > \pm d) d\mathbf{x}'.
\end{aligned}
$$

Now applying the same rotation ρ_0 as in the proof of (i) to the integral variable $\mathbf{x}'$, since Φ is radial, the value of the functional is

$$
\begin{aligned}
& < a^{1-n/2}(\phi_M^\rho)_a, l > \\
= & \hat{\Phi}(\mathbf{0}) a^{1-n} \int_{\mathbf{R}^n} e^{i < \rho_0^{-1}\rho^{-1}\mathbf{k}, \frac{\mathbf{x}'}{a} >} \Phi(\frac{\mathbf{x}'}{a}) \\
& \delta(< \rho_0^{-1}\mathbf{a}, \mathbf{x}' > \pm d) d\mathbf{x}' \\
= & \hat{\Phi}(\mathbf{0}) a^{1-n} \int_{\mathbf{R}^n} e^{i < \rho_0^{-1}\rho^{-1}\mathbf{k}, \frac{\mathbf{x}'}{a} >} \Phi(\frac{\mathbf{x}'}{a}) \delta(x_n' \pm d) d\mathbf{x}' \\
= & \hat{\Phi}(\mathbf{0}) a^{1-n} \int_{\mathbf{R}^{n-1}} e^{i < \rho_0^{-1}\rho^{-1}\mathbf{k}, \frac{(x_1', \cdots, x_{n-1}', \mp d)}{a} >} \\
& \Phi(\frac{(x_1', \cdots, x_{n-1}', d)}{a}) dx_1' \cdots dx_{n-1}' \\
= & \hat{\Phi}(\mathbf{0}) \int_{\mathbf{R}^{n-1}} e^{i < \rho_0^{-1}\rho^{-1}\mathbf{k}, (x_1', \cdots, x_{n-1}', \mp d/a) >} \\
& \Phi((x_1', \cdots, x_{n-1}', d/a)) dx_1' \cdots dx_{n-1}' \\
= & \hat{\Phi}(\mathbf{0}) e^{\mp i(d/a)(\rho_0^{-1}\rho^{-1}\mathbf{k})_n} \\
& \mathcal{F}^{(n-1)}(\Phi_{(d/a)})((\rho_0^{-1}\rho^{-1}\mathbf{k})^{(n-1)}),
\end{aligned}
$$

where $\mathcal{F}^{(n-1)}$ denotes the Fourier transform in $\mathbf{R}^{n-1}$ for $(x_1', \cdots, x_{n-1}')$. Denote the j-th component of a general vector $\mathbf{v}$ by $(\mathbf{v})_j$, and $(\mathbf{v})^{(n-1)} = ((\mathbf{v})_1, \cdots, (\mathbf{v})_{n-1})$. In this notation,

$$
\rho_0^{-1}\rho^{-1}\mathbf{k} = ((\rho_0^{-1}\rho^{-1}\mathbf{k})^{(n-1)}, (\rho_0^{-1}\rho^{-1}\mathbf{k})_n).
$$

Set

$$
\Phi_{(\mathbf{v})_\mathbf{n}}((\mathbf{v})^{(n-1)}) = \Phi(\mathbf{v}).
$$

It is noted that $\Phi_{(\mathbf{v})_n}$ is a radial function in $(v)^{(n-1)}$. Let θ be the angle between the vector $\rho_0^{-1}\rho^{-1}\mathbf{k}$ and the vector $(0,\cdots,0,1)$ that is identical with the angle between the vector $\rho^{-1}\mathbf{k}$ and the normal of the hyperplane $<\mathbf{a},\mathbf{x}>=0$. We have

$$\begin{aligned}(\rho_0^{-1}\rho^{-1}\mathbf{k})_n &= <\rho_0^{-1}\rho^{-1}\mathbf{k},(0,\cdots,0,1)> \\ &= (\cos\theta)|\rho_0^{-1}\rho^{-1}\mathbf{k}| = (\cos\theta)|\mathbf{k}|,\end{aligned}$$

and therefore

$$|(\rho_0^{-1}\rho^{-1}\mathbf{k})^{(n-1)}| = (\sin\theta)|\mathbf{k}|.$$

Using this notation, the last value of the functional can be further written

$$\hat{\Phi}(\mathbf{0})e^{\pm i(\cos\theta)|\mathbf{k}|d/a}\psi_{(d/a)}((\sin\theta)|\mathbf{k}|),$$

where we adopt the notation

$$\mathcal{F}^{(n-1)}(\Phi_{(\mathbf{v})_n})((\xi)^{(n-1)}) = \psi_{(\mathbf{v})_n}(|(\xi)^{(n-1)}|).$$

When we prove the assertion (ii) there are two things are different from the above: The constant multiple in ϕ_E is $\hat{\Phi}(-\mathbf{k})$ in stead of $\hat{\Phi}(\mathbf{0})$ in ϕ_M; and in the integral representing the functional the exponential function is missing. These induce the final result $\hat{\Phi}(\mathbf{k})\psi_{(d/a)}(0)$. The proof is complete.

Note that the assumption that Φ is radial in Theorem 7.4 is only to make the functions appearing in the end of the formulas for assertions (ii) and (iii) to be radial and thus simpler. There is a general version of the theorem for Φ not being a radial function.

By applying Theorem 6.4 to the special case $\Phi(\mathbf{x}) = e^{-(1/2)|x|^2}$, we obtain

Corollary 7.1. *For* $\Phi(\mathbf{x}) = e^{-(1/2)|x|^2}, \mathbf{x}\in\mathbf{R}^n$, *we have*

(i) *For any* $a>0$ *and* $\mathbf{b}\in\mathbf{R}^n$, *in the pointwise sense,*

$$\begin{aligned}&a^{n/2-1}(W^{(2)}_{\phi_E}l)(a,\rho,\mathbf{b})\\ &= (2\pi)^{(n-1)/2}e^{-(1/2)|\mathbf{k}|^2}e^{-d^2/(2a^2)},\end{aligned}$$

where d *is the distance from* $\mathbf{b}$ *to the hyperplane* $<\mathbf{a},\mathbf{x}>=0$.

(ii) *For any* $a>0$ *and* $\mathbf{b}\in\mathbf{R}^n$, *in the pointwise sense,*

$$\begin{aligned}&a^{1-\frac{n}{2}}(W^{(2)}_{\phi_M}l)(a,\rho,\mathbf{b})\\ =\ &(2\pi)^{\frac{n-1}{2}}e^{i(\cos\theta)|\mathbf{k}|d/a}e^{-\frac{d^2}{2a^2}}e^{-\frac{1}{2}\sin^2\theta|\mathbf{k}|^2}.\end{aligned}$$

For $n=2$ these are (ii) and (iii) of Theorem 4.1. The assertion (i) of Theorem 4.1 is from (i) of Theorem 7.4. Theorem 7.4 and Corollary 7.1 well establish the applicability of rotational wavelets to edge detection. Briefly speaking, based on these results, ϕ_M^ρ may be used to detect the portions of a $(n-1)$-dimensional surface whose normal vectors are close to the direction of the vector $\rho^{-1}\mathbf{k}$. It is expected that a large class of functions Φ satisfy the conditions in Theorem 6.4 for which the Fourier transform function $\psi_{(d/a)}$ may be explicitly worked out. Whenever $\psi_{(d/a)}(\sin\theta)|\mathbf{k}|)$ fast decays along with $d/a\to\infty$ and takes its maximal value at $\theta=0$ for $d=0$ (Owing to Riemann–Lebesgue Theorem, in any case, $\theta\neq 0, \psi_{(d/a)}(\sin\theta|\mathbf{k}|)\to 0$ as $|\mathbf{k}|\to\infty$), then we can use the function ϕ_M to detect edge. We especially note that for the edge detection purpose only (i), (ii) and (iv) are essential. We, in particular, do not need the admissibility condition (4) as we do not use inversion formula. As example, if $P(\mathbf{x})$ is a polynomial of degree k and $\alpha>0$, then Φ can be of the following forms as long as the moment is non-zero:

$$P(\mathbf{x})e^{-\alpha|\mathbf{x}|^2},\quad P(\mathbf{x})e^{-\alpha|\mathbf{x}|},\quad \text{and}\quad \frac{P(\mathbf{x})}{(1+|\mathbf{x}|)^{n+k+\alpha}}.$$

Acknowledgment. The authors would like to thank Professor Tang Ping for her various kind support that makes the collaboration on this work and this paper possible.

REFERENCES

[1] F.Brackx and F.Sommen, Clifford-Hermite Wavelets in Euclidean Space, *The Journal of Fourier Analysis and Application*. Vol.6. pp.299-310 (2000).

[2] S.Mallat, *A Wavelet Tour of Signal Processing*, Academic Press. San Diego (1999).

[3] J.P.Antoine and P.Vandergheynst, Two-Dimensional Directional Wavelets in Image Processing, *International Journal of Imaging Systems and Technology,* Vol.7. pp.152-165 (1999).

[4] J.P. Antoine, R.Murenzi and B.Piette, Image Analysis with $2D$ Continuous Wavelet Transform: Detection of Position, Orientation and Visual Contrast of Simple Objects, *Proceedings of the International Conference,* France, pp.144-159 (1989).

[5] J.Li, *Wavelet Analysis and Signal Processing – Theory, Applications and Software Implementations,* Chongqing Publishing House, China (1997).

[6] Y.Peng, *Wavelet Transform and Applications in Engineering,* Science Publishing House, China (1999).

[7] A.K.Louis, P. Maab and A. Rieder, *Wavelet Theory And Applications,* John Wiley & Sons Press, Chichester (1997).

[8] J.P.Antoine, *The continuous wavelet transform in image processing,* CWI Quarterly 11 (1998), pp. 1-23.

[9] J.P.Antoine, *Wavelet analysis of signals and images: A grand tour,* Revista Ciencias Matematicas (La Habana) 18 (2000) pp. 113-143.

[10] C.K.Chui, *An Introduction to Wavelets,* Academic Press, Inc., San Diego, CA (1992).

[11] E.Hewitt and K.Stromberg, *Real and Abstract Analysis,* Springer-Verlag, Berlin (1999).

[12] E.M.Stein and G.Weiss, *Introduction to Fourier Analysis on Euclidean Spaces,* Princeton University Press, Princeton, NJ (1999).

[13] L.Zhang and P.Lenders, *Morlet Wavelet Based Human Head and Face Boundary Extraction,* Proceedings of Fifth Asian Conference on Computer Vision, Melbourne (2002),pp 801-806.

[14] L.Zhang and P.Lenders, Locating the Head Boundary with 2D Continuous Wavelet Transform, *Proceedings of 2001 International Symposium on Intelligent Multimedia, vedio & Speech Processing*, Hong Kong (2001), pp336-339.

[15] L.Zhang and P.Lenders, Human Head and Face Boundary Detection Based on Region Shield Segmentation and Wavelet Transform, *Proceedings of the 6th International Conference on Control, Automation, Robotics and Vision*, Singapore (2000).

[16] T. Qian and L. Zhang, Radon Measure Model For Edge Detection Using Rotational Wavelets, *Proceedings of WSEAS International Conference on Wavelet Analysis and Multirate Systems,* December 2002, Greece.

[17] J. Canny, 1986, A computational approach to edge detection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol.PAMI8, pp. 679698.

E-mail address: LMZhang@umac.mo

E-mail address: fsttq@umac.mo

E-mail address: qingyezeng@hotmail.com