Pattern Recognition Letters, 2012

Reliable Ear Identification using 2-D Quadrature Filters

Tak-Shing Chan, Ajay Kumar

This paper investigates a new approach for reliable personal identification using gray-level ear images. The developed approach extracts robust phase information using 2-D quadrature filtering (both monogenic and quaternionic). The use of quadrature filters is motivated by their ability to jointly localize spatial and frequency domain phase information in the segmented ear images. This paper has developed two new techniques for biometric recognition: the quaternionic quadrature filter and the QuaternionicCode. We comparatively evaluate the performance of 1-D and 2-D local phase information using Cauchy, Gaussian derivative and log-Gabor band-pass quadrature filters. We detail experimental results, both for recognition and verification, using publicly available UND ear dataset and IITD ear database. The achieved results with 2-D quadrature filters are highly promising and achieve significantly improved performance as compared to conventional phase encoding using 1-D quadrature filters employed in the literature.

1. Introduction

Automated personal identification using unique physiological characteristics has invited significant research and development efforts in the literature. Among various biometric modalities, personal identification using human face imaging has the highest user acceptance and therefore has attracted the majority of research efforts in biometrics. Current face recognition algorithms are highly sensitive to pose, illumination, expression and ageing. The face profile images often contain ear images which can be exploited to improve the limiting performance of face recognition.

The human ear is a highly stable structure which is rich in information and easy to be imaged for biometric identification. The ear shape is also invariant to or least influenced by facial expressions and can also be filmed covertly for surveillance applications. The ear imaging is also highly immune to privacy, stigma, and hygienic issues associated with the collection of traditional biometrics (Bertillon, 1890; Hurley *et al.*, 2005; Kumar & Zhang, 2007). Unlike face, fingerprint, iris or other hand-based biometric modalities, the ear has invited little attention in the biometrics literature.

Recent efforts to develop remote biometric systems are inviting renewed interests in reliable ear identification. In order to make a case for the commercial deployment of ear-based personal identification systems, the performance of currently available algorithms for ear segmentation and matching needs to be significantly improved.

The uniqueness of human ears has been extensively studied. Earlier studies by Iannarelli (1989) on over 10,000 ears samples suggested the uniqueness of ear topologies as each ear in his sample was unique based on twelve measurements. Burge & Burger (2000) characterized the ear using Voronoi diagrams but did not go further to report systematic experimental results on ear databases. Chang *et al.* (2003) exploited principal component analysis (PCA) to characterize the face and the ear. They showed that each of these two modalities has comparable performances for personal identification, and that the combination of both can achieve significant improvement in recognition performance. Hurley *et al*. (2005) presented a novel approach to ear identification called the force field transform. They achieved a rank-one recognition rate of 99.2% (compared to 62.4% for PCA) on the XM2VTS dataset, but on a limited probe dataset of 63 subjects. Using a sparse representation framework on a smaller UND subset (32 subjects), Naseem *et al*. (2008) achieved an equal error rate of 5% while their rank-one recognition rate was 91.67%. Fourier descriptors can be exploited to develop rotation-invariant representations of ear images and such an approach has been detailed in Abate *et al.* (2006). The encoding of ear profile data representing shape and wrinkles for neural-network-based ear identification has been explored by Moreno *et al.* (1999). The shape and structural features can also be extracted from the gradient of ear images and such a characterization has been attempted in Mu *et al.* (2004). Bustard & Nixon (2010) have recently exploited the scale-invariant feature transform (SIFT) for matching ear images and illustrated promising results. The literature review

on 2-D ear identification suggests that researchers have exploited a range of spectral (Abate *et al.*, 2006), structural (Chang *et al.*, 2003; Moreno *et al.* 1999; Mu *et al.*, 2004; Hurley *et al.*, 2002; Hurley *et al.*, 2005; Choras, 2004; Abdel-Mottaleb *et al.*, 2006; Arbab-Zavar & Nixon, 2007) and 3-D (Chen & Bhanu, 2007; Bhanu & Chen, 2003; Theoharis *et al.*, 2008; Yan & Bowyer, 2007) features for the characterization of ear images. However such approaches have only been tested on small scale databases and in most cases the size of the probe dataset has been less than 100 subjects. In addition, the use of varying number of subjects, databases, and the importantly evaluation protocols (for the use of same database in rare cases) makes it very difficult to ascertain the best approach that can make a convincing case for the development of ear-based biometric identification for commercial applications.

1.1 Our Work

This paper presents a new approach for personal identification using ear images. The developed approach encodes reliable phase information using 2-D quadrature filtering. We extensively evaluate both quaternionic and monogenic quadtature filters and develop a new QuaternionicCode based approach for the ear identification. Our experimental results suggests that the performance from the quaternionic quadrature filters and the monogenic quadtature filters consistently outperform 1D log-Gabor filter based approach presented in the literature. The 1-D log-Gabor filter based matching approach is used as baseline as it has shown to outperform several other competing approaches on publicly available ear databases.

The development of personal identification systems using the ear biometric requires robust and automated segmentation of region of interest images from the face profile images acquired in a contactless but cooperative environment. In this work, we employed gray-level ear images acquired using contactless ear imaging while the

automated segmentation of ear (region of interest) was achieved using the approach based on morphological operators and Fourier descriptors as detailed in Kumar & Wu (2011). The enhanced segmented ear images are subjected to a new feature extraction approach using 2-D quadrature filtering which is followed by phase quantization to generate template images for the matching. We present extensive experimental results to ascertain the superiority of 2-D quadrature filters both for the verification and recognition problems.

This paper is organized as follows. Section 2 contains the theoretical foundations for 2-D analytic signals, followed by Section 3 which describes feature extraction and matching using quadrature filters and Hamming distances. Section 4 reports experimental results. We conclude with discussion and conclusions in Sections 5 and 6 respectively.

2. Theoretical Foundation and Background

2.1 Quaternion Algebra

The algebra of quaternions (Hamilton, 1844) is an extension of the 2-D complex algebra to four dimensions and is defined as:

$$q = w + ix + jy + kz, \tag{2.1}$$

where $w, x, y, z \in \mathbf{R}$ and i, j, k are imaginary units such that $i^2 = j^2 = k^2 = ijk = -1$. If $w = 0$, q is called *pure*. From this definition, it readily follows that:

$$ij = k, jk = i, ki = j, ji = -k, kj = -i, ik = -j. \tag{2.2}$$

Thus quaternion multiplication is not commutative. The conjugate, norm and inverse of a quaternion are defined, respectively, as:

$$\bar{q} = w - ix - jy - kz, \tag{2.3}$$

$$|q| = \sqrt{w^2 + x^2 + y^2 + z^2}, \tag{2.4}$$

$$q^{-1} = \frac{\bar{q}}{|q|^2}. \tag{2.5}$$

If $|q| = 1$, it is called a *unit quaternion*.

2.2 Quaternion Fourier Transform

The quaternion Fourier transform (QFT) and inverse quaternion Fourier transform (IQFT) can be defined in three different ways, where the complex exponentials ($e^{\pm\mu_1 2\pi ux}$ and $e^{\pm\mu_2 2\pi vy}$) appear on (i) both sides, (ii) the left side, and (iii) the right side, respectively (Ell, 1993; Pei *et al.*, 2001). As quaternion multiplication is non-commutative, these definitions are not equivalent. The following definitions closely follow those given in Pei *et al.* (2001).

Two-sided QFT and IQFT:

$$F_1(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\mu_1 2\pi ux} f(x,y) e^{-\mu_2 2\pi vy} dxdy \tag{2.6}$$

$$f(x,y) = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{\mu_1 2\pi ux} F_1(u,v) e^{\mu_2 vy} dudv \tag{2.7}$$

Left-sided QFT and IQFT:

$$F_2(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\mu_1 2\pi(ux+vy)} f(x,y) dxdy \tag{2.8}$$

$$f(x,y) = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{\mu_1 2\pi(ux+vy)} F_2(u,v) dudv \tag{2.9}$$

Right-sided QFT and IQFT:

$$F_3(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) e^{-\mu_1 2\pi(ux+vy)} dxdy \tag{2.10}$$

$$f(x,y) = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F_3(u,v) e^{\mu_1 2\pi(ux+vy)} dudv \tag{2.11}$$

Here μ_1, μ_2 are pure, unit quaternions. The above transforms can be discretized in the

obvious manner (Pei *et al.*, 2001).

Two-sided QFT and IQFT:

$$F_1(u,v) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} e^{-\mu_1 2\pi\frac{ux}{M}} f(x,y) e^{-\mu_2 2\pi\frac{vy}{N}} \tag{2.12}$$

$$f(u,v) = \frac{1}{MN}\sum_{u=0}^{M-1}\sum_{v=0}^{N-1} e^{\mu_1 2\pi\frac{ux}{M}} F_1(x,y) e^{\mu_2 2\pi\frac{vy}{N}} \tag{2.13}$$

Left-sided QFT and IQFT:

$$F_2(u,v) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} e^{-\mu_1 2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)} f(x,y) \tag{2.14}$$

$$f(u,v) = \frac{1}{MN}\sum_{u=0}^{M-1}\sum_{v=0}^{N-1} e^{\mu_1 2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)} F_2(x,y) \tag{2.15}$$

Right-sided QFT and IQFT:

$$F_3(u,v) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y) e^{-\mu_1 2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)} \tag{2.16}$$

$$f(u,v) = \frac{1}{MN}\sum_{u=0}^{M-1}\sum_{v=0}^{N-1} F_3(x,y) e^{\mu_1 2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)} \tag{2.17}$$

In the rest of this paper, we will use the two-sided QFT and IQFT with $\mu_1 = i$ and $\mu_2 = j$, because this coincides with the definition used by Bülow & Sommer (2001) who proposed the quaternionic analytic signal we used below. Another justification for the two-sided QFT is based on the convolution theorem. The convolution theorem enables us to perform spatial domain filtering via multiplication in the frequency domain, which can be time-saving (González & Woods, 2008). Bülow *et al.* (2001) proved that if both $g(x,y)$ and the two-sided QFT of $h(x,y)$ are real-valued, then the convolution theorem holds:

$$f(x,y) = (g * h)(x,y) \overset{F_1}{\Leftrightarrow} F(u,v) = G(u,v)H(u,v). \tag{2.18}$$

For left-sided QFT, Pei *et al.* (2001) proved that if $h(x,y) = h(-x,-y)$, then we have:

$$f(x,y) = (g * h)(x,y) \overset{F_2}{\Leftrightarrow} F(u,v) = G(u,v)H(u,v). \tag{2.19}$$

Finally, for right-sided QFT, Pei *et al.* (2001) also proved that if $g(x,y) = g(-x,-y)$, then:

$$f(x,y) = (g * h)(x,y) \overset{F_3}{\Leftrightarrow} F(u,v) = G(u,v)H(u,v). \tag{2.20}$$

In other cases, the convolution theorem does not hold and we must resort to more complicated expressions in the frequency domain (Pei *et al.*, 2001). Since the quaternionic filter is real-valued in the frequency domain (cf. Section 3) and that the input image is also real, the two-sided QFT is the best choice for our purposes (Bülow *et al.*, 2001).

2.3 Analytic, Quaternionic and Monogenic Signals

For many application domains such as pattern recognition and biometrics, the local phase of a signal is of great importance. The local phase of a 1-D real signal can be obtained from its analytic signal (Boukerroui *et al.*, 2004). The analytic signal of a real function $f(x)$ is defined by (Hahn, 1996):

$$f_A(x) = f(x) + i\left(\frac{1}{\pi x} * f(x)\right), \tag{2.21}$$

with $1/\pi x$ denoting the Hilbert kernel. In the Fourier domain this can be written as:

$$F_A(u) = \big(1 + \operatorname{sgn}(u)\big)F(u), \tag{2.22}$$

where

$$\operatorname{sgn}(u) = \begin{cases} 1, & u > 0 \\ 0, & u = 0 \\ -1, & u < 0. \end{cases} \tag{2.23}$$

A few attempts have been made to extend the analytic signal to two dimensions, for example, the partial analytic signal, the total analytic signal, the quaternionic signal and the monogenic signal (Boukerroui *et al.*, 2004). The most recent ones are the quaternionic and monogenic signals. The quaternionic signal is defined by Bülow

(1999) in the frequency domain as:

$$F_Q(u,v) = \big(1 + \mathrm{sgn}(u)\big)\big(1 + \mathrm{sgn}(v)\big)F_1(u,v), \tag{2.24}$$

which extends the analytic signal to the quaternionic Fourier domain. Unfortunately, the quaternionic signal is not isotropic (Felsberg & Sommer, 2001; Boukerroui *et al.*, 2004), meaning that it is not robust to changes in orientation. An isotropic alternative, the monogenic signal, is obtained by replacing the Hilbert kernel in the original definition (f_A) with the 2-D Riesz kernel, making it a vector-valued analytic signal (Felsberg & Sommer, 2000, 2001):

$$f_M(x,y) = \big(f(x,y), (h_1 * f)(x,y), (h_2 * f)(x,y)\big) \tag{2.25}$$

where

$$\big(h_1(x,y), h_2(x,y)\big) = \left(\mp\frac{x}{2\pi(x^2+y^2)^{3/2}}, \mp\frac{y}{2\pi(x^2+y^2)^{3/2}}\right), \tag{2.26}$$

noticing that the signs are negative in Felsberg & Sommer (2000) but positive in Felsberg & Sommer (2001). This kernel can be written in the Fourier domain as:

$$\big(H_1(u,v), H_2(u,v)\big) = \left(\pm i\frac{u}{\sqrt{u^2+v^2}}, \pm i\frac{v}{\sqrt{u^2+v^2}}\right), \tag{2.27}$$

ditto about the signs (Felsberg & Sommer, 2000, 2001). It is worth mentioning that the monogenic signal can be embedded into the quaternionic Fourier domain via the following mapping (Felsberg & Sommer, 2001):

$$F_M(u,v) = \frac{\sqrt{u^2+v^2} \mp (1,k)(u,v)^T}{\sqrt{u^2+v^2}} F(u,v), \tag{2.28}$$

but we will not pursue this further. For this paper, we will adopt the earliest definition of the monogenic signal (Felsberg & Sommer, 2000), which is also the one implemented by Kovesi (2000).

3. Feature Extraction and Matching

3.1 Quadrature Filters

The analytic signal is not appropriate for applications requiring both space and frequency localizations (Boukerroui *et al.*, 2004). In order to circumvent this limitation,

we need quadrature filters which compute the analytic signal of a band-pass filtered signal (Boukerroui *et al.*, 2004). In what follows, we will extend three 1-D quadrature filters from Boukerroui *et al.* (2004) to their 2-D counterparts. Here W_{-1} and W_0 denote two branches of the Lambert W function which is the multivalued inverse function of xe^x (Euler, 1783; Corless *et al.*, 1996).

The 2-D Cauchy filter:

$$H_B(u,v) = (u^2+v^2)^{a/2} \exp\left(-\sigma\sqrt{u^2+v^2}\right) \tag{3.1}$$

where a/σ is the center frequency and a is the bandwidth parameter. The bandwidth of this filter is given by (Boukerroui *et al.*, 2004):

$$b = \log_2 \frac{W_{-1}(c)}{W_0(c)}, c = -\frac{2^{-1/a}}{e}. \tag{3.2}$$

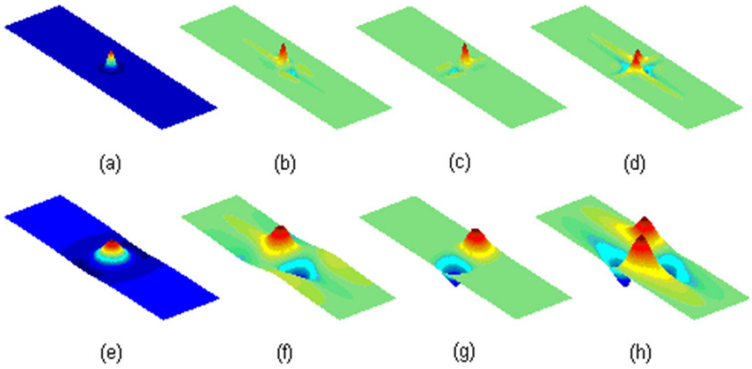


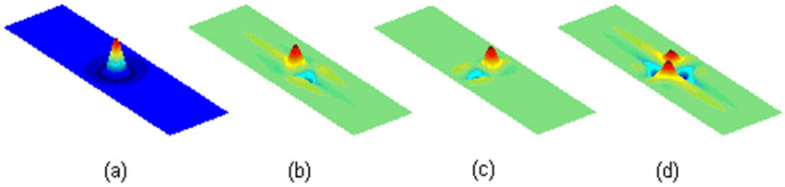
Figure 1: Spatial representation of *quaternionic Cauchy filters* having a bandwidth of 2.0310 octaves. Top row: (a) real and (b-d) $\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}$ parts with a center frequency of 1/18. Bottom row: (e) real and (f-h) $\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}$ parts with a center frequency of 1/54.

The 2-D Gaussian derivative filter:

$$H_B(u,v) = (u^2+v^2)^{a/2} \exp\left(-\sigma^2(u^2+v^2)\right) \tag{3.3}$$

where $\sqrt{a}/\sigma$ is the center frequency and a is the bandwidth parameter. The bandwidth of this filter is given by (*op. cit.*):

$$b = \frac{1}{2}\log_2 \frac{W_{-1}(c)}{W_0(c)}, c = -\frac{4^{-1/a}}{e}. \tag{3.4}$$



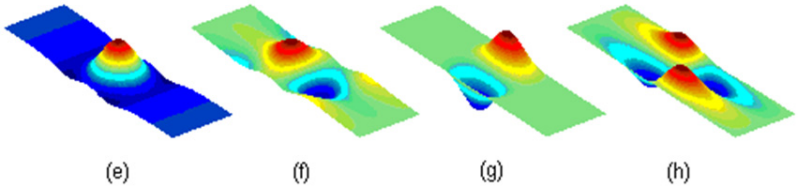


Figure 2: Spatial representation of *quaternionic Gaussian derivative* filters with a bandwidth of 2.0310 octaves. Top row: (a) real and (b-d) $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ parts with a center frequency of 1/18. Bottom row: (e) real and (f-h) $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ parts with a center frequency of 1/54.

The 2-D log-Gabor filter (cf. Felsberg & Sommer, 2000):

$$H_B(u,v) = \exp\left(-\frac{\ln^2\left(\sqrt{u^2+v^2}/f_0\right)}{2\ln^2(\sigma/f_0)}\right) \tag{3.5}$$

where f_0 is the center frequency and σ/f_0 is the bandwidth parameter. The bandwidth of this filter is given by (*op. cit.*):

$$b = -2\sqrt{2\ln 2}\log_2(\sigma/f_0). \tag{3.6}$$

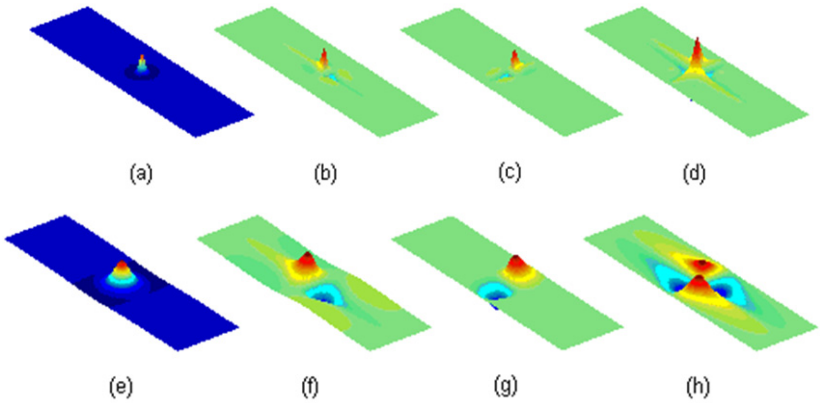
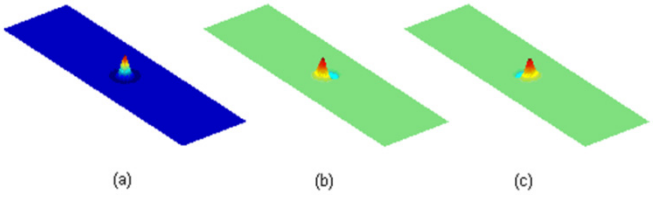


Figure 3: Spatial representation of *quaternionic log-Gabor* filters having a bandwidth of 2.0310 octaves. Top row: (a) real and (b-d) $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ parts with a center frequency of 1/18. Bottom row: (e) real and (f-h) $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ parts with a center frequency of 1/54.

Now we define the quaternionic quadrature filter (QQF) as the quaternionic signal of a 2-D band-pass filter H_B:

$$QQF = \big(1+\mathrm{sgn}(u)\big)\big(1+\mathrm{sgn}(v)\big)F_1(u,v)H_B \tag{3.7}$$



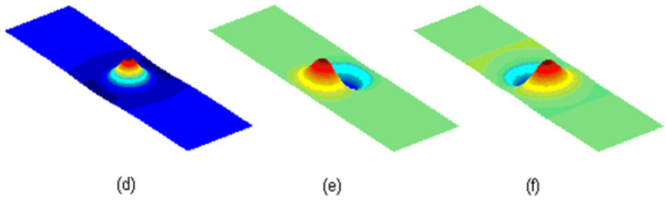


Figure 4: Spatial representation of *monogenic Cauchy filters* having a bandwidth of 2.0310 octaves. Top row: (a) real and (b-c) vector parts with a center frequency of 1/18. Bottom row: (d) real and (e-f) vector parts with a center frequency of 1/54.

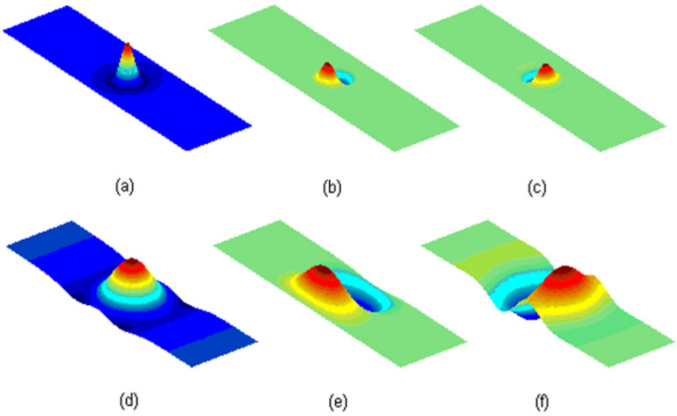


Figure 5: Spatial representation of *monogenic Gaussian derivative* filters having a bandwidth of 2.0310 octaves. Top row: (a) real and (b-c) vector parts with a center frequency of 1/18. Bottom row: (d) real and (e-f) vector parts with a center frequency of 1/54.

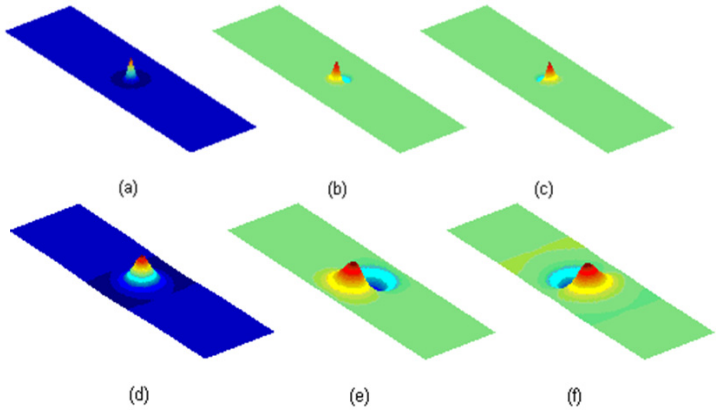


Figure 6: Spatial representation of *monogenic log-Gabor filters* having a bandwidth of 2.0310 octaves. Top row: (a) real and (b-c) vector parts with a center frequency of 1/18. Bottom row: (d) real and (e-f) vector parts with a center frequency of 1/54.

Note that our definition differs from the original quaternionic Gabor filter which uses a band-pass filter alone to approximate the quaternionic signal (Bülow & Sommer, 1998; Bülow & Sommer, 2001). Their limitation is that the band-pass filter must reside in the

first quadrant of the frequency domain or else it would cease to produce a quaternionic signal; *our new formulation does not suffer from this restriction*. Finally,

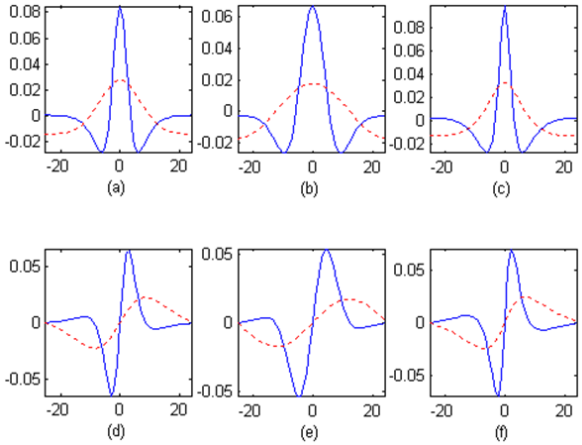


Figure 7: Spatial representation of 1-D filters having a bandwidth of 2.0310 octaves and center frequencies of 1/18 (solid) and 1/54 (dotted). Top row: real part of (a) Cauchy, (b) Gaussian derivative and (c) log-Gabor filters. Bottom row: imaginary part of (d) Cauchy, (e) Gaussian derivative and (f) log-Gabor filters.

monogenic quadrature filters, or spherical quadrature filters (SQF), are defined as the monogenic signal of a 2-D band-pass filter H_B (Felsberg & Sommer, 2000):

$$SQF = (H_B, H_1 H_B, H_2 H_B). \tag{3.8}$$

3.2 Phase Quantization and Matching

Daugman (2003) defined phase quantization by:

$$A_{\{Re,Im\}} = \mathrm{sgn}^{+}_{\{Re,Im\}} \iint_{r\theta} f(r,\theta) h(r,\theta; r_0, \theta_0, \alpha, \beta, \omega) r dr d\theta \tag{3.9}$$

where $\mathrm{sgn}^{+}_{\{Re,Im\}}$ extracts the nonnegative signs of the real and imaginary parts of a complex number, respectively, $f(r,\theta)$ is an iris image and $h(r,\theta; r_0, \theta_0, \alpha, \beta, \omega)$ is a 2-D Gabor wavelet (both in polar coordinates). Daugman (2003) further defined the normalized Hamming distance between the codes A and B (with masks A_M and B_M) as:

$$D(A,B) = \frac{\|(A \otimes B) \cap A_M \cap B_M\|}{\|A_M \cap B_M\|}, \tag{3.10}$$

where $\cap$ is the bitwise AND operator and $\otimes$ is the bitwise exclusive OR operator.

The monogenic quadrature filters (Koveski, 2000; Felsberg & Sommer, 2000) can also be combined with Dougman's (2003) phase quantization strategy to generate binarized templates as follows:

$$A_1(x,y) = \text{sgn}^+\big((h_b * f)(x,y)\big), \tag{3.11}$$

$$A_2(x,y) = \text{sgn}^+\big((h_1 * h_b * f)(x,y)\big), \tag{3.12}$$

$$A_3(x,y) = \text{sgn}^+\big((h_2 * h_b * f)(x,y)\big). \tag{3.13}$$

The template size of above binarized image representation is three bits per pixels which can be referred to as the *MonogenicCode* (Zhang *et al.*, 2010). The distance between such binarized image templates can also be computed using the normalized Hamming distance as follows (without masks):

$$D(A,B) = \frac{\sum_{y=1}^{N}\sum_{x=1}^{M}\sum_{i=1}^{3} A_i(x,y) \otimes B_i(x,y)}{3MN}. \tag{3.14}$$

In this work, we explore a new image representation which is referred to *QuaternionicCode*. The *QuaternionicCode* is essentially binarized phase representation of an image using quaternionic filters and computed as follows:

$$A_{\{1,2,3,4\}}(x,y) = \text{sgn}^+_{\{1,\text{i},\text{j},\text{k}\}}\big((qqf * f)(x,y)\big), \tag{3.15}$$

where $\text{sgn}^+_{\{1,\text{i},\text{j},\text{k}\}}$ extracts the nonnegative signs of the $1, i, j, k$ parts of a quaternion, respectively, qqf is the spatial domain representation of QQF defined in (3.7), and $f(x,y)$ is an input image. Our template size is four bits per pixel. The normalized Hamming distance can be modified accordingly:

$$D(A,B) = \frac{\sum_{y=1}^{N}\sum_{x=1}^{M}\sum_{i=1}^{4} A_i(x,y) \otimes B_i(x,y)}{4MN}. \tag{3.16}$$

The block diagram of completely automated ear identification using *QuaternionicCode* representation is shown in Figure 8. The steps of automated ear segmentation, are exactly same as detailed in (Kumar and Wu, 2011). The segmented images are unsharpened and employed for feature extraction using quadrature filters. The quadrature filtered images are used to generate binarized templates, using equation 3.15, which are employed to generate matching scores.

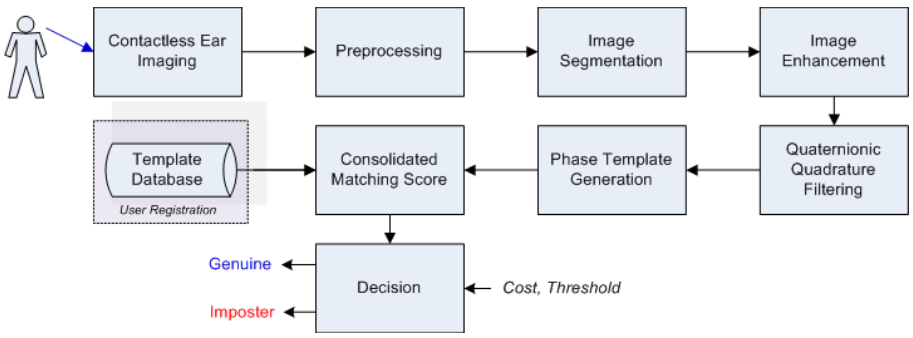


Figure 8: Block diagram of human identification system using ear images.

4. Experiments and Results

In this section we report experimental results from the 2-D quadrature filter-based approach detailed in previous section. We performed extensive experiments with one training and two training image samples on two publicly available ear image datasets. Methods and results are detailed in following subsections.

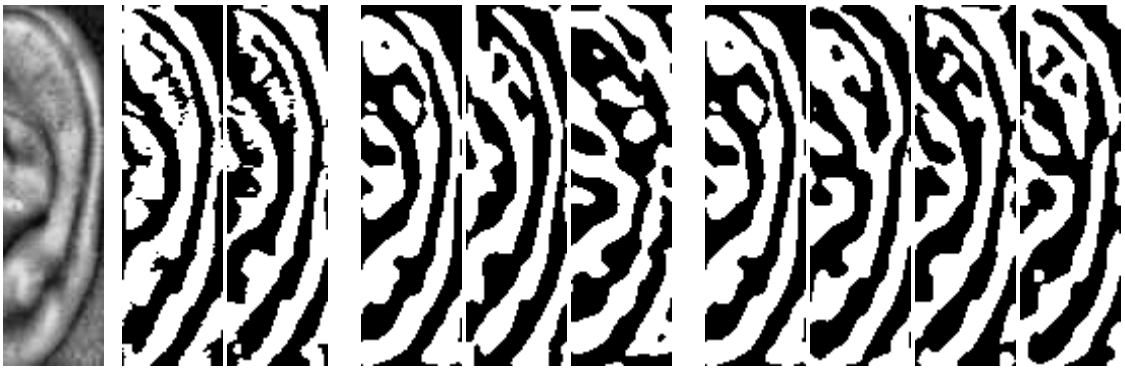
4.1 Evaluation Datasets

The method developed in this paper was evaluated on two publicly available ear image datasets. The first dataset is the UND dataset (Chang *et al.*, 2003) which contains images from 110 subjects. The second dataset is the IITD ear image database (Kumar & Wu, 2011; IITD, 2011) which contains ear images from 125 subjects. The IITD database also provides a larger dataset of ear images from 221 subjects, which is formed by the combination of the 125-subject dataset with another dataset. The steps of automated ear segmentation and enhancement are exactly same as detailed in (Kumar & Wu, 2011) for both UND and IITD dataset. This approach utilizes combinations of morphological operations to probe and coarsely locate the ear shape boundary. These boundaries refined using Fourier descriptors to generate ear shape profile and identify two stable points. These stable points correspond to two points on the refined ear shape profile that have maximum distance among them. Each of the ear images are subjected to rotational and scale normalization and segment region of interest. These segmented

images are further subjected to illumination normalization and the resulting grayscale images of size 50×180 pixels are employed in this work. Each of the two databases was also enlarged to include {0, 3, -3, 6, -6, 9, -9} degrees of rotated samples (using bicubic interpolation) to accommodate rotational variations in the ear images. Figure 9 shows typical image samples from the two dataset employed in this evaluation. Figure 10 shows typical samples of the encoded templates generated using monogenic and quaternionic encoding.

(a) (b)

Figure 9: Ear image samples from publicly available (a) IITD and (b) UND dataset.



(a) (b) (c) (d)

Figure 10: Ear processed with original and extended log-Gabor filters having a bandwidth of 2.0310 octaves and center frequency of 1/18: (a) enhanced image, (b) 1-D encoding {re, im}, (c) monogenic encoding {1, h1, h2} and (d) quaternionic encoding {1, i, j, k}. Note that the real part of (c) and (d) are identical, as expected.

4.2 Verification Experiments

The key objective of verification experiments is to ascertain the comparative performance of 1D and 2D quadrature filters for the ear verification problem. In all verification experiments we compare each subject to every other subject on a

one-to-one basis using the normalized Hamming distance measure. In order to ascertain the performance of the proposed approach using minimum or one training image and two training image, we report experimental results from two test protocols. The test protocol A generates average of three tests where each of the first three images from each of the subjects are used as test images while remaining other respective images are used as training images. The test protocol B (all-to-all) reports average of performance when every image of the every subject in the dataset is used as respective training image/sample. In each of the verification experiments, we report average of experimental results using equal error rate (EER) and receiver operating characteristics (ROC). For the 125 subject IITD dataset, using protocol A with three tests, we generate 125×3 genuine comparisons and 124×125×3 imposter comparisons, and we generate 493 genuine comparisons and 124×493 imposter comparisons using protocol B. For the 221 IITD dataset, using protocol A with three tests, we generate 221×3 genuine comparisons and 220×221×3 imposter comparisons while 793 genuine comparisons and 220×793 imposter comparisons are generated using protocol B. Finally, for the UND dataset, we generate 110×3 genuine comparisons and 109×110×3 imposter comparisons using protocol A, and 433 genuine comparisons and 109×443 imposter comparisons using protocol B.

We experimented with quaternionic and monogenic filters on three sets of ear data (125, 221 and UND), with 1-D quadrature filters as a baseline. Specifically, Cauchy, Gaussian derivative, and log-Gabor filters are used, with bandwidths of 2.0310 octaves and wavelengths of 18 and 54 (these filter parameters are same as used by Kumar & Wu, 2011 for log-Gabor filters and generate the best performance). The filters are plotted in Figures 2 to 8. Verification results are summarized in Tables 1-3 (best results in blue) while the corresponding receiver operating characteristics are shown in figure 11-13.

Table 1: Equal error rates for quaternionic quadrature filters.

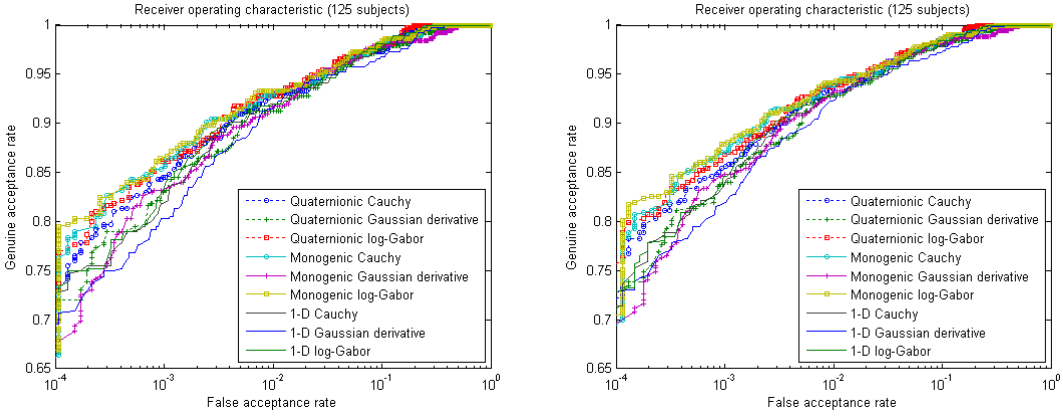
	Protocol A			Protocol B		
	125	221	UND	125	221	UND
Quaternionic Cauchy	4.21	2.71	8.13	3.65	2.65	7.67
Quaternionic Gaussian derivative	4.31	3.47	6.36	3.85	3.40	6.37
Quaternionic log-Gabor	3.73	2.84	7.27	3.78	2.66	7.16

Table 2: Equal error rates for monogenic quadrature filters.

	Protocol A			Protocol B		
	125	221	UND	125	221	UND
Monogenic Cauchy	4.05	2.41	6.97	3.65	2.29	6.68
Monogenic Gaussian derivative	4.01	2.98	6.10	3.65	2.90	6.00
Monogenic log-Gabor	4.00	2.41	6.97	3.66	2.38	6.47

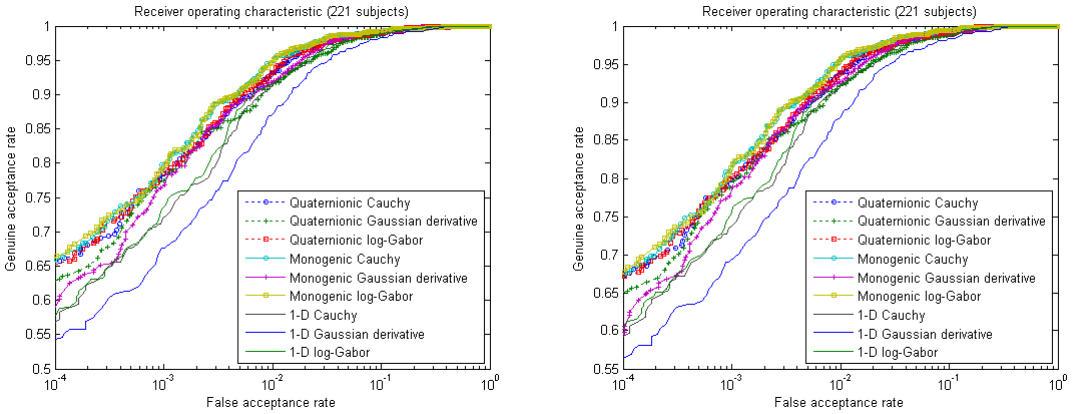
Table 3: Equal error rates for 1-D quadrature filters.

	Protocol A			Protocol B		
	125	221	UND	125	221	UND
Cauchy	4.27	3.32	7.58	3.90	3.15	7.13
Gaussian derivative	4.75	4.07	7.62	4.09	3.78	7.16
Log-Gabor	4.32	3.13	8.18	4.06	3.13	7.40



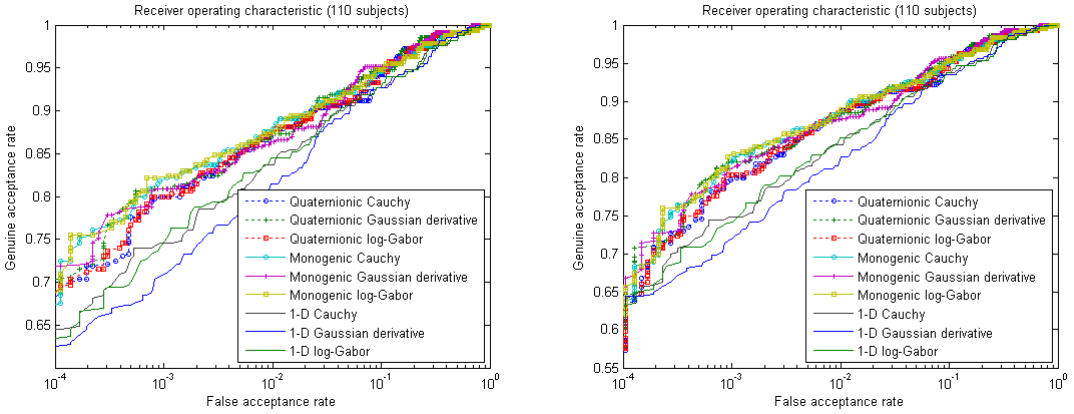
(a) (b)

Figure 11: ROC for the 125 database: (a) protocol A and (b) protocol B.



(a) (b)

Figure 12: ROC for the 221 database: (a) protocol A and (b) protocol B.



(a) (b)

Figure 13: ROC for the UND database: (a) protocol A and (b) protocol B.

The experimental results for the verification using 125 subject ear dataset with protocol A suggest that the best performing 2-D and 1-D quadrature filters are quaternionic log-Gabor (EER = 3.73%) and 1-D Cauchy (EER = 4.27%) respectively. For the 221 subject dataset, the best performance is achieved from monogenic Cauchy/monogenic log-Gabor (EER = 2.41%) and 1-D log-Gabor (EER = 3.13%). For the UND dataset, the best ones are monogenic Gaussian derivative (EER = 6.10%) and 1-D Cauchy (EER = 7.58%). The best performing filters for 125 subject dataset using protocol B are quaternionic Cauchy/monogenic Cauchy/monogenic Gaussian derivative (EER = 3.65%) and 1-D Cauchy (EER = 3.90%). Corresponding best performing filters for 221 subject dataset are monogenic Cauchy (EER = 2.29%), and

1-D log-Gabor (EER = 3.13%). For the UND dataset, the best filters are monogenic Gaussian derivative (EER = 6.00%) and 1-D Cauchy (EER = 7.13%). With the exception of the 125 subject dataset, the best performing filters are robust to changes in the protocol. In other words the best performing filter for one dataset is likely to be best again when we switch the test protocols. Another key observation is that the 2-D quadrature filters outperformed the 1-D filters in the verification experiments. It is worth noting that there has not been any effort in the literature to explore ear verification or recognition performance using monogenic log-Gabor (or quadrature) filters. Our experimental results (table 2-3, also table 5-6) suggests that monogenic log-Gabor filters consistently outperforms conventional 1D log Gabor filters based approach presented in the literature.

4.3 Recognition Experiments

In all the recognition experiments, we follow the same two protocols as discussed in Subsection 4.2, using the same quadrature filters on the same three sets of data. The test protocol A generates average recognition accuracy from three tests where each of 3 images from each of the subjects are used as probe image while remaining other respective images are used as gallery images. The test protocol B generates average of the recognition performance when every image of the every subject is used as the gallery while other respective images of that subject are used as the probe images. The average rank-one recognition rate (R1RR) is reported from all the experiments in the Tables 4 to 6 (best results in blue). We also illustrate the cumulative match characteristic (CMC) curves from each of the recognition experiments in the figure 14-16.

The recognition experiments on 125 subject dataset using protocol A suggest that the best performing 2-D and 1-D quadrature filters are quaternionic

Cauchy/quaternionic log-Gabor (R1RR = 96.53%) and 1-D log-Gabor (R1RR = 95.73%) respectively. Similarly for the 221 subject dataset, the best filters are monogenic log-Gabor (R1RR = 96.08%) and 1-D log-Gabor (R1RR = 94.72%). For

Table 4: Average rank-one recognition rates for quaternionic quadrature filters.

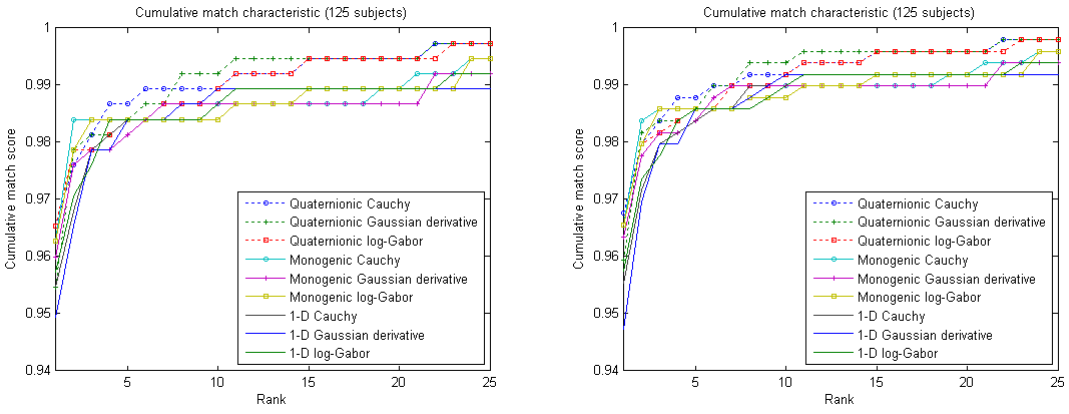
	Protocol A			Protocol B		
	125	221	UND	125	221	UND
Quaternionic Cauchy	96.53	95.17	90.30	96.75	95.33	90.99
Quaternionic Gaussian derivative	95.47	94.87	91.82	95.94	95.33	92.38
Quaternionic log-Gabor	96.53	95.32	90.30	96.55	95.46	90.99

Table 5: Average rank-one recognition rates for monogenic quadrature filters.

	Protocol A			Protocol B		
	125	221	UND	125	221	UND
Monogenic Cauchy	96.27	95.93	91.52	96.55	96.09	91.92
Monogenic Gaussian derivative	96.00	94.72	91.52	96.35	95.08	91.92
Monogenic log-Gabor	96.27	96.08	91.21	96.55	96.34	91.69

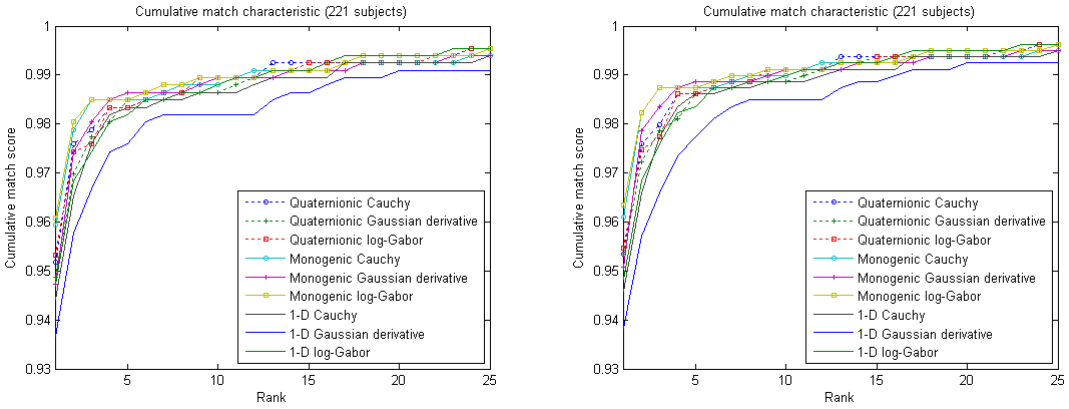
Table 6: Average rank-one recognition rates for 1-D quadrature filters.

	Protocol A			Protocol B		
	125	221	UND	125	221	UND
Cauchy	95.47	94.42	89.09	95.54	94.58	89.61
Gaussian derivative	94.93	93.67	87.27	94.73	93.82	88.45
Log-Gabor	95.73	94.72	87.88	95.74	94.83	88.91



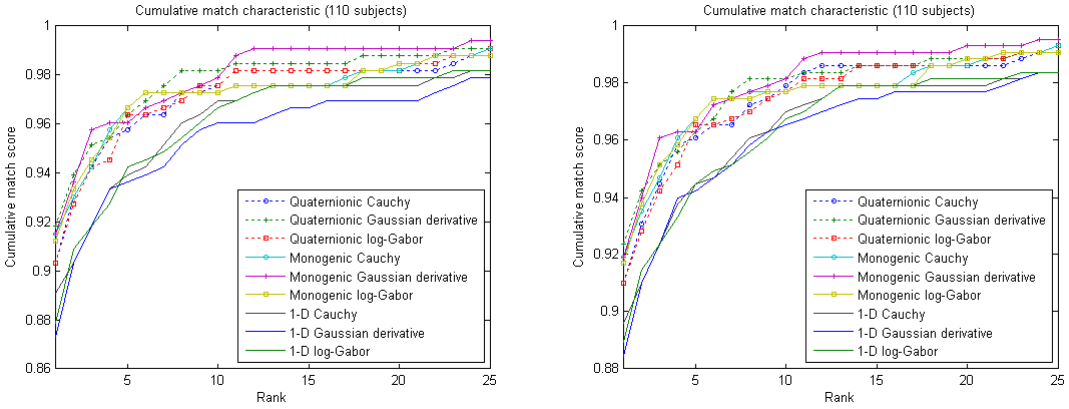
(a) (b)

Figure 14: CMC for the 125 database: (a) protocol A and (b) protocol B.



(a) (b)

Figure 15: CMC for the 221 database: (a) protocol A and (b) protocol B.



(a) (b)

Figure 16: CMC for the UND database: (a) protocol A and (b) protocol B.

the UND dataset, the best ones are quaternionic Gaussian derivative (R1RR = 91.82%) and 1-D Cauchy (R1RR = 89.09%). The protocol B for 125 subject dataset delivers best performance from quaternionic Cauchy (R1RR = 96.75%) and 1-D log-Gabor (R1RR = 95.74%). For the 221 subject dataset, the best performing filters are monogenic log-Gabor (R1RR = 96.34%) and 1-D log-Gabor (R1RR = 94.83%). For the UND dataset, the best ones are quaternionic Gaussian derivative (R1RR = 92.38%) and 1-D Cauchy (R1RR = 89.61%). Again the best performing filters are robust against changes in the test protocols. Our experimental results for both recognition and verification experiments also suggest that the 2-D quadrature filters achieves better performance than corresponding 1-D quadrature filters.

5. Discussion

There has not been any effort in the literature to exploit monogenic or any other forms of 2-D quadrature filters for ear identification while log-Gabor 1-D filter-based approach has shown its superiority over Gabor phase, Gabor orientation, force field transform, and geometrical feature-based approaches on the publicly available ear dataset (Kumar & Wu, 2011). Therefore we employed the 1-D log-Gabor filter-based approach as a baseline and evaluated its performance against the quaternionic and monogenic quadrature filter-based approaches developed in this paper. The experimental results from verification and recognition experiments consistently suggest the superior performance of monogenic log-Gabor filter over 1-D log-Gabor filter. Experimental results presented in the previous section also suggest that 2-D quadrature filters consistently outperformed their 1-D counterparts.

Initially, we expected monogenic filters to work best in every case because of its isotropic property, but it turned out that although MonogenicCode was best on average, QuaternionicCode worked better in a several cases (*e.g.*, EER and R1RR for the 125 dataset for both the test protocol A and protocol B). Clearly, the no free lunch theorem (Wolpert & Macready, 1997) applies here, and given this, it might be a good idea to explore/develop the alternative formulations of 2-D quadrature filters even if the resulting filters are theoretically suboptimal. One possibility is to look at other hypercomplex representations for signal processing (Alfsmann, 2007).

Our verification results (ROC) disagreed with Boukerroui *et al.* (2004) who claimed that Cauchy filters are better than log-Gabor for feature detection. This can be possibly because our two-octave bandwidth may be too low for Cauchy filters to work best. In general, the best verification results are achieved by using monogenic log-Gabor filters but closely followed by monogenic Cauchy filters. However, this

observation is not consistent for the recognition results where the quaternionic quadrature filters are observed to achieve superior performance. The goals for the recognition and the verification problem are different (Bolle *et al*., 2005). Therefore any specific approach or the matcher which performs better for the verification problem may not be the best for recognition problem. Our experimental results have consistently achieved superior recognition results using quaternionic quadrature filters as compared to those using conventional 1-D log-Gabor filter based approach and is therefore recommended for ear recognition.

6. Conclusions

This paper has developed a new approach using 2-D quadrature filters, including monogenic and quaternionic filters, for automated ear identification. We proposed a new quaternionic quadrature filter which is more general than Bülow & Sommer's (1998) original quaternionic Gabor filter. The proposed filter is of significance since it enables us to use arbitrary band-pass filters that are not necessarily shifted to the positive frequency quadrant. We also proposed new approach for biometric feature representation using QuaternionicCode and evaluated its performance for the ear identification. Our experimental results in this paper, presented on the two publicly available databases, have demonstrated the superiority of the 2-D quadrature filter approach.

Our efforts to exploit 2-D quadrature for the biometric identification have achieved significant results in this paper. However, further efforts are required to ascertain the effectiveness of phase information extracted from the 2-D quaternionic quadrature filters for other biometric modalities. We have not yet compared the quaternionic quadrature filter with the original quaternionic Gabor filter. It is part of our future work since it would be interesting to ascertain whether their 2-D shift to the

positive frequency quadrant would affect the ear recognition and verification performances. Further work should also be directed to explore the effectiveness of directional filters (Granlund & Knutsson, 1995) for the biometrics identification.

References

Ianarelli, A. (1989). *Ear Identification*. Fremont, CA: Paramount Publishing Company.

Hamilton, W. R. (1844). On quaternions, or on a new system of imaginaries in algebra. *Philosophical Magazine*, 25:10—13.

Ell, T. A. (1993). Quaternion-Fourier transforms for analysis of two-dimensional linear time-invariant partial differential systems. In *Proceedings of the 32nd IEEE Conference on Decision and Control* (pp. 1830—1841). Piscataway, NJ: IEEE Press.

Pei, S.-C., Ding, J.-J. and Chang, J.-H. (2001). Efficient implementation of quaternion Fourier transform, convolution, and correlation by 2-D complex FFT. *IEEE Transactions on Signal Processing*, 49(11):2783—2797.

Bülow, T. and Sommer, G. (2001). Local hypercomplex signal representations and applications. In *Geometric Computing with Clifford Algebras* (pp. 255—289). London: Springer-Verlag.

González, R. C. and Woods, R. E. (2008). *Digital Image Processing*. Upper Saddle River, NJ: Prentice Hall.

Bülow, T., Felsberg, M. and Sommer, G. (2001). Non-commutative hypercomplex Fourier tranforms of multidimensional signals. In *Geometric Computing with Clifford Algebras* (pp. 187—207). London: Springer-Verlag.

Boukerroui, D., Noble, J. A. and Brady, M. (2004). On the choice of band-pass quadrature filters. *Journal of Mathematical Imaging and Vision*, 21:53—80.

Hahn, S. L. (1996). *Hilbert Transforms in Signal Processing*. Boston: Artech House.

Bülow, T. (1999). *Hypercomplex spectral signal representations for the processing and analysis of images*. Unpublished PhD thesis, University of Kiel, Kiel, Germany.

Felsberg, M. and Sommer, G. (2000). A new extension of linear signal processing for estimating local properties and detecting features. In *Proceedings of the 22nd DAGM Symposium für Mustererkennung* (pp. 195—202). Berlin: Springer-Verlag.

Felsberg, M. and Sommer, G. (2001). The monogenic signal. *IEEE Transactions on Signal Processing*, 49(12):3136—3144.

Kovesi, P. D. (2005). *MATLAB and Octave Functions for Computer Vision and Image Processing*. Matching image phase points using monogenic phase data, Available from:

<http://www.csse.uwa.edu.au/~pk/research/matlabfns/Match/matchbymonogenicphase.m >.

Euler, L. (1783). De serie Lambertina, plurimisque eius insignibus proprietatibus. *Acta Academiae Scientarum Imperialis Petropolitinae*, 2:29—51.

Corless, R., Gonnet, G., Hare, D., Jeffrey, D. and Knuth, D. (1996). On the Lambert W function. *Advances in Computational Mathematics*, 5:329—359.

Zhang, L., Zhang, L. and Zhang, D. (2010). MonogenicCode: A novel fast feature coding algorithm with applications to finger-knuckle-print recognition. In *Proceedings of the 2010 International Workshop on Emerging Techniques and Challenges for Hand-Based Biometrics* (pp. 1—4). Piscataway, NJ: IEEE Press.

Bülow, T. and Sommer, G. (1998). Quaternionic Gabor filters for local structure classification. In *Proceedings of the Fourteenth International Conference on Pattern Recognition* (pp. 808—810). Piscataway, NJ: IEEE Press.

Dougman, J. (2003). The importance of being random: Statistical principles of iris recognition. *Pattern Recognition*, 36(2):279—291.

Kumar, A. and Wu, C. (2012). Automated human identification using ear imaging. *Pattern Recognition*, 41(5), 2012.

Chang, K., Bowyer, K. W., Sarkar, S. and Victor, B. (2003). Comparison and combination of ear and face images in appearance-based biometrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(9):1160—1165.

Wolpert, D. H., Macready, W. G. (1997). No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*, 1(1):67—82.

Hurley, D. J., Nixon, M. S., and Carter, J. N. (2005). Force field feature extraction for ear biometrics. *Computer Vision and Image Understanding*, 98(3):491—512.

Naseem, I., Togneri, R. and Bennamoun, M. (2008). Sparse representation for ear biometrics. In G. Bebis *et al.* (Eds.), *Advances in Visual Computing* (pp. 336—345). Berlin: Springer-Verlag.

Alfsmann, D., Gockler, H. G., Sangwine, S. J. and Ell, T. A. (2007). Hypercomplex algebras in digital signal processing: Benefits and drawbacks. In *Proceedings of the 15th European Signal Processing Conference* (pp. 1322—1326). European Association for Signal Processing.

Granlund, G. H. and Knutsson, H. (1995). *Signal Processing for Computer Vision*. Dordrecht: Kluwer Academic Publishers.

Burge, M. and Burger, W. (1998). Ear biometrics. In A. K. Jain, R. Bolle and S. Pankanti (Eds.), *Biometrics: Personal Identification in Networked Society* (pp. 273—286). Boston, MA: Kluwer Academic Publishers.

Hurley, D. J., Nixon, M. S. and Carter, J. N. (2002). Force field energy functionals for image feature

extraction. *Image and Vision Computing*, 20(5/6):311—317.

Moreno, B., Sanchez, A. and Velez, J. F. (1999). On the use of outer ear images for personal identification in security applications. In *Proceedings of the IEEE 33rd Annual 1999 International Carnahan Conference on Security Technology* (pp. 469—476). Piscataway, NJ: IEEE Press.

Bhanu, B. and Chen, H. (2003). Human ear recognition in 3D. In *Proceedings of the 2003 Workshop on Multimodal User Authentication* (pp. 91—98).

Chen, H. and Bhanu, B. (2007). Human ear recognition in 3D. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(4):718—737.

Mu, Z., Yuan, L., Xu, Z., Xi, D. and Qi, S. (2004). Shape and structural feature based ear recognition. In S. Li *et al.* (Eds.), *Advances in Biometric Person Authentication* (pp. 311—364). Berlin: Springer-Verlag.

Arbab-Zavar, B., Nixon, M. S. and Hurley, D. J. (2007). On model-based analysis of ear biometrics. In *Proceedings of the First IEEE International Conference on Biometrics: Theory, Applications, and Systems* (pp. 1—5). Piscataway, NJ: IEEE Press.

Yan, P. and Bowyer, K. W. (2007). Biometric recognition using 3D ear shape. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(8):1297—1308.

Choras, M. (2004). Ear biometrics based on geometrical method of feature extraction. In F. Perales and B. Draper (Eds.), *Articulated Motion and Deformable Objects* (pp. 51—61). Berlin: Springer-Verlag.

Abate, A. F., Nappi, M., Riccio, D. and Ricciardi, S. (2006). Ear recognition by means of a rotation invariant descriptor. In *Proceedings of the 18th International Conference on Pattern Recognition* (pp. 437—440). Piscataway, NJ: IEEE Press.

IITD (2010). *IIT Delhi Ear Database Version 1*. Available from: <http://webold.iitd.ac.in/~biometrics/Database_Ear.htm>.

Kumar, A. and Zhang, D. (2007). Ear authentication using log-Gabor wavelets. *Proceedings of the SPIE*, 6539:65390A.

Adbel-Mottaleb, M. and Zhou, J. (2006). Human ear recognition from face profile images. In D. Zhang and A. Jain (Eds.), *Advances in Biometrics* (pp. 786—792). Berlin: Springer-Verlag.

Bustard, J. D. and Nixon, M. S. (2010). Towards unconstrained ear recognition from two-dimensional images. *IEEE Transactions on Systems, Man and Cybernetics A*, 40(3):486—494.

Theoharis, T., Passalis, G., Toderici, G. and Kakadiaris, I. (2008). Unified 3D face and ear recognition using wavelets on geometry images. *Pattern Recognition*, 41(3):796—804.

Bolle, R.M. Connell, J.H. Pankanti, S. Ratha, N.K. Senior, A.W (2005), The relation between ROC curve and CMC, *Proc. 4th IEEE Workshop on Automatic Identification Advanced Technologies* (pp 15-20).