Computer Science and Artificial Intelligence Laboratory
Technical Report

MIT-CSAIL-TR-2014-009 April 26, 2014

Quaternionic Representation of the Riesz Pyramid for Video Magnification

Neal Wadhwa, Michael Rubinstein, Fredo Durand, and William T. Freeman

massachusetts institute of technology, cambridge, ma 02139 usa — www.csail.mit.edu

Quaternionic Representation of the Riesz Pyramid for Video Magnification

Neal Wadhwa[1] Michael Rubinstein[1,2] Frédo Durand[1] William T. Freeman[1]
[1]MIT CSAIL [2]Microsoft Research
http://people.csail.mit.edu/nwadhwa/riesz-pyramid

Abstract

Recently, we presented a new image pyramid, called the Riesz pyramid, that uses the Riesz transform to manipulate the phase in non-oriented sub-bands of an image sequence to produce real-time motion-magnified videos. In this report we give a quaternionic formulation of the Riesz pyramid, and show how several seemingly heuristic choices in how to use the Riesz transform for phase-based video magnification fall out of this formulation in a natural and principled way. We intend this report to accompany the original paper on the Riesz pyramid for video magnification.

1. Introduction

Numerous phenomena exhibit small motions that are invisible to the naked eye. These motions require computational amplification to be revealed [3, 4, 6, 8]. Manipulating the local phase in coefficients of a complex steerable pyramid decomposition of an image sequence is an effective way to amplify small motions. However, complex steerable pyramids are very overcomplete (21 times) and costly to construct, making them unsuitable for real-time processing.

Recently, we presented a new image pyramid representation, the Riesz pyramid, that is suitable for real-time Eulerian phase-based video processing [7]. This representation consists of first decomposing an image into non-oriented sub-bands and then applying an approximate Riesz transform, the natural two dimensional generalization of the Hilbert transform, to each sub-band. The Riesz transform is a two channel filter bank, so the coefficients of the Riesz pyramid are triples consisting of the sub-band value and two Riesz transform values [1, 5]. We showed that the Riesz pyramid can be used for phase-based motion magnification because the Riesz transform is a steerable Hilbert transformer that allows us to compute a quadrature pair that is 90 degrees out of phase with the original sub-band with respect to the dominant orientation at every pixel. This means we only phase-shift and translate image features in the direction of the dominant orientation rather than in a sampling of orientations as done in the complex steerable pyramid. This vastly reduces the overcompleteness of the representation and makes the computation a lot more efficient.

However, this formulation has an ambiguity as the local phase at a point can be positive or negative depending on whether the orientation is expressed by an angle θ or its antipode $\theta + \pi$. This ambiguity is not present in the complex steerable pyramid as the orientation of each sub-band is fixed. This causes problems when performing spatiotemporal filtering on the phases, especially in regions where the orientation angle wraps around regardless of what specific angle is specified as the wraparound point (Fig. 1(c)).

We resolve this problem by turning to the quaternionic formulation of the Riesz pyramid triples [1]. Instead of using the local phase as specified before, we spatiotemporally filter the *quaternionic phase*, which is invariant to the aforementioned ambiguity. This quaternionic phase is related to a quaternion by the quaternion logarithm in the same way the phase of a complex number is related to it by the complex logarithm.

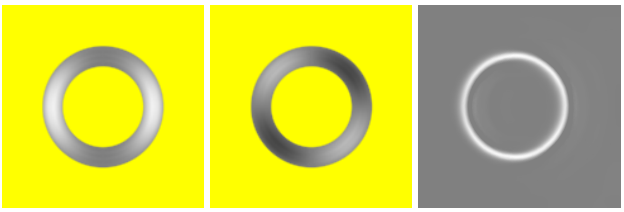
This work is meant to supplement Wadhwa et al. 2014 [7], where we used the quaternionic phase we derive in this report without mathematical justification. This report concentrates on the technical details and definitions of how to use quaternions to perform phase analysis and video magnification. Further details about the Riesz pyramid and motion magnification results on natural and synthetic sequences can be found in [7].

2. Background

A Riesz pyramid coefficient consists of an input subband I and two Riesz transform values (R_1, R_2). This triple can be used to determine the local amplitude A, local phase ϕ and local orientation θ [5]. In Eq. 3 in Wadhwa et al. 2014 [7], reproduced here, we related the Riesz pyramid coefficient with these three quantities via the equation

$$I = A\cos(\phi), R_1 = A\sin(\phi)\cos(\theta), R_2 = A\sin(\phi)\sin(\theta) \quad (1)$$

(a) Input (b) Phase difference (c) Amplified with (b)



(d) $\phi \cos(\theta)$ difference (e) $\phi \sin(\theta)$ difference (f) Amplified with (d,e)

Figure 1. The motion between the input (a) and a copy shifted to the left by one half pixel is magnified in two ways. First, the phase difference of ϕ (b) is spatially denoised and then used to magnify the second frame (c). In the bottom row, the difference in the quantities $\phi\cos(\theta)$ and $\phi\sin(\theta)$ (d-e) are spatially denoised and then used to amplify the second frame (f). In (b,d,e), low amplitude regions are masked in yellow, middle gray corresponds to a difference of zero and only a single sub-band is shown.

That is, the local phase ϕ and orientation θ are angles in a spherical representation of the value (I, R_1, R_2). While we can solve for these angles, there is not a unique solution. If (A, ϕ, θ) is a solution to Eq. 1, so is $(A, -\phi, \theta + \pi)$. Our solution to this problem was to instead consider

$$\phi\cos(\theta), \phi\sin(\theta) \tag{2}$$

which are invariant to this sign ambiguity. If the Riesz pyramid coefficient is viewed as a quaternion, then Eq. 2 is the quaternion logarithm of the normalized coefficient. This is completely analogous to the complex steerable pyramid, in which the phase is the complex logarithm of a complex coefficient. This justifies our use of Eq. 2 as a proxy for the motion. We review complex exponentiation and logarithms and quaternions in the remainder of this section.

Complex Exponential and Logarithm The exponential function of a complex number $z = a + ib$ can be defined through the power series

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!} = e^a(\cos(b) + i\sin(b)). \tag{3}$$

The inverse (logarithm) of the map is not well-defined as multiple complex inputs map to the same output. For example, $e^{i2\pi} = e^{i4\pi} = 1$, which means that $\log(1)$ could equal $i2\pi$ or $i4\pi$ in addition to its usual value of 0. This problem can be solved by restricting the imaginary part of the logarithm to lie in the range $[-\pi, \pi)$. That is, the imaginary part takes its *principal value*. Then the logarithm of a complex number $z = a + ib$ is

$$\log(z) = \log(\sqrt{a^2 + b^2}) + i\text{atan2}(b, a). \tag{4}$$

In the case of a unit complex number, the first term is zero and we are left with

$$\log(z) = i\text{atan2}(b, a), \tag{5}$$

which is just the principle value of the phase of the complex number. Therefore, complex logarithm and exponentiation give a useful way to go from a complex number to its phase and back again.

Quaternions Quaternions are a generalization of the complex numbers, in which there are three imaginary units, denoted i, j and k, so that each quaternion is characterized by four numbers, one real and three imaginary. Quaternion multiplication is associative and distributive with addition and can therefore be fully defined by the following property of the imaginary units:

$$-1 = i^2 = j^2 = k^2 = ijk. \tag{6}$$

Specifically, multiplication is given by

$$\begin{array}{rl} (q_1 + iq_2 + jq_3 + kq_4)(r_1 + ir_2 + jr_3 + k4_r) & = \\ (q_1r_1 - q_2r_2 - q_3r_3 - q_4r_4) & + \\ i(q_1r_2 + q_2r_1 + q_3r_4 - q_4r_3) & + \\ j(q_1r_3 - q_2r_4 + q_3r_1 + q_4r_2) & + \\ k(q_1r_4 + q_2r_3 - q_3r_2 + q_4r_1). & \end{array} \tag{7}$$

Note that multiplication is noncommutative.

For a quaternion $\mathbf{q} = q_1 + iq_2 + jq_3 + kq_4$, its conjugate $\mathbf{q}^*$, norm $\|\mathbf{q}\|$ and inverse $\mathbf{q}^{-1}$ are defined as

$$\mathbf{q}^* = q_1 - iq_2 - jq_3 - kq_4, \tag{8}$$

$$\|\mathbf{q}\| = \sqrt{q_1^2 + q_2^2 + q_3^2 + q_4^2}, \tag{9}$$

$$\mathbf{q}^{-1} = \mathbf{q}^*/\|\mathbf{q}\|^2, \tag{10}$$

where the third definition follows from the first two. The exponential of a quaternion $\mathbf{q} = q_1 + \mathbf{v}$ (where $\mathbf{v} = iq_2 + jq_3 + kq_4$) is defined by its power series as with the complex number (Eq. 3),

$$e^{\mathbf{q}} = \sum_{n=0}^{\infty} \frac{\mathbf{q}^n}{n!} = e^{q_1}\left(\cos(\|\mathbf{v}\|) + \frac{\mathbf{v}}{\|\mathbf{v}\|}\sin(\|\mathbf{v}\|)\right). \tag{11}$$

The inverse of this function is

$$\log(\mathbf{q}) = \log(\|\mathbf{q}\|) + \frac{\mathbf{v}}{\|\mathbf{v}\|}\text{acos}\left(\frac{q_1}{\|\mathbf{q}\|}\right). \tag{12}$$

In the case of a unit quaternion, where $\|\mathbf{q}\| = 1$, this simplifies to

$$\frac{\mathbf{v}}{\|\mathbf{v}\|}\text{acos}(q_1). \tag{13}$$

This is an imaginary quantity that is analogous to Eq. 5. We refer to it as the *quaternionic phase* to distinguish it from the local phase in the non-quaternionic formulation of the Riesz pyramid in [7].

3. Quaternion Representation of Riesz Pyramid

Now that we have defined quaternions, we can represent the Riesz pyramid coefficient triple (I, R_1, R_2) as a quaternion $\mathbf{r}$ with the original subband I being the real part and the two Riesz transform components (R_1, R_2) being the imaginary i and j components of the quaternion

$$\mathbf{r} = I + iR_1 + jR_2 \tag{14}$$

or if we use Eq. 1, we can write this as

$$\mathbf{r} = A\cos(\phi) + iA\sin(\phi)\cos(\theta) + jA\sin(\phi)\sin(\theta). \tag{15}$$

Rather then solving for the local amplitude A, phase ϕ and orientation θ, we instead use the quaternionic phase we defined earlier. That is, we can express the amplitude and quaternionic phase (Fig. 2(c)) as

$$\begin{aligned} A = &\quad \|\mathbf{r}\| \\ i\phi\cos(\theta) + j\phi\sin(\theta) = &\quad \log(\mathbf{r}/\|\mathbf{r}\|) \end{aligned} \tag{16}$$

The second quantity is computed by applying Eq. 13 to the specific case of normalized Riesz pyramid coefficients and is invariant to whether the local phase and orientation are ϕ and θ or the antipode $-\phi$ and $\theta + \pi$.

4. Video Magnification with Quaternions

We perform video magnification in a similar way as Wadhwa et al. 2013 (Fig. 2). However, rather than computing the local phase, which has an ambiguous sign, we instead compute the quaternionic phase of the Riesz pyramid coefficients and filter them first in time and then in space. This resolves the sign issue and is completely analogous to Wadhwa et al. 2013, except instead of filtering the complex logarithm of the normalized coefficients in a complex steerable pyramid, we filter the quaternionic logarithm of the normalized coefficients in a Riesz pyramid.

4.1. Filtering of Quaternionic Phase

While the quaternionic phase resolves the sign ambiguity in filtering $\phi(x, y, t)$ directly, it is still a wrapped quantity. That is, $i\phi\cos(\theta) + j\phi\sin(\theta)$ and $i(\phi + 2\pi)\cos(\theta) + j(\phi + 2\pi)\sin(\theta)$ correspond to the same value. Therefore, instead of filtering the quaternionic phase directly, we instead use a technique by Lee and Shin [2] to filter a sequence of unit quaternions. This technique is tantamount to phase unwrapping the quaternionic phases in time and then performing LTI filtering. We use it to LTI filter the Riesz pyramid coefficients at each pixel in each scale in time and then in a subsequent step we spatially smooth the pixel values with an amplitude weighted blur to improve SNR. We will also make the assumption that the local orientation at any pixel is roughly constant in time and approximately locally constant in space.

Suppose at a single location (x, y) in a single scale ω_r, the normalized Riesz pyramid coefficients are

$$\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n \tag{17}$$

where $\mathbf{r}_m = \cos(\phi_m) + i\sin(\phi_m)\cos(\theta_m) + j\sin(\phi_m)\sin(\theta_m)$, the most general form of a unit quaternion with no k component.

In ordinary complex phase unwrapping, we would take the principle value of the difference between successive terms and then do a cumulative sum to give an unwrapped sequence in which the difference between two successive terms was always in the interval $(-\pi, \pi]$. We do the same thing here. We compute the principle value of the phase difference between successive coefficients by dividing them and then taking the logarithm.

$$\log(\mathbf{r}_1), \log(\mathbf{r}_2\mathbf{r}_1^{-1}), \ldots, \log(\mathbf{r}_n\mathbf{r}_{n-1}^{-1}) \tag{18}$$

The terms $\mathbf{r}_m\mathbf{r}_{m-1}^{-1}$ will in general have nonzero k component (Eq. 7). However, if we make the assumption that $\theta_m = \theta + \epsilon$, that is that the local orientation is roughly constant over time at every pixel, the k term will be close to zero. More explicitly,

$$\begin{aligned} \mathbf{r}_m\mathbf{r}_{m-1}^{-1} = &\quad \cos(\phi_m - \phi_{m-1}) \\ + &\quad i\sin(\phi_m - \phi_{m-1})\cos(\theta) \\ + &\quad j\sin(\phi_m - \phi_{m-1})\sin(\theta) + O(\epsilon) \end{aligned} \tag{19}$$

which, ignoring the $O(\epsilon)$ term, has logarithm

$$i([\phi_m - \phi_{m-1}])\cos(\theta) + j([\phi_m - \phi_{m-1}])\sin(\theta) \tag{20}$$

where the bracketed terms are taken modulo 2π.

The second step is perform a cumulative sum of Eq. 18

$$\phi_1\mathbf{u}, (\phi_1 + [\phi_2 - \phi_1])\mathbf{u}, \ldots, \left(\phi_1 + \sum_{l=2}^{n}[\phi_l - \phi_{l-1}]\right)\mathbf{u} \tag{21}$$

where $\mathbf{u} = i\cos(\theta) + j\sin(\theta)$. If we let $\phi'_m = \phi_1 + \sum_{l=2}^{m}[\phi_l - \phi_{l-1}]$, we can more compactly write this series as

$$i\phi'_m\cos(\theta) + j\phi'_m\sin(\theta) \tag{22}$$

At every pixel, we perform temporal filtering on this quantity to isolate motions of interest.

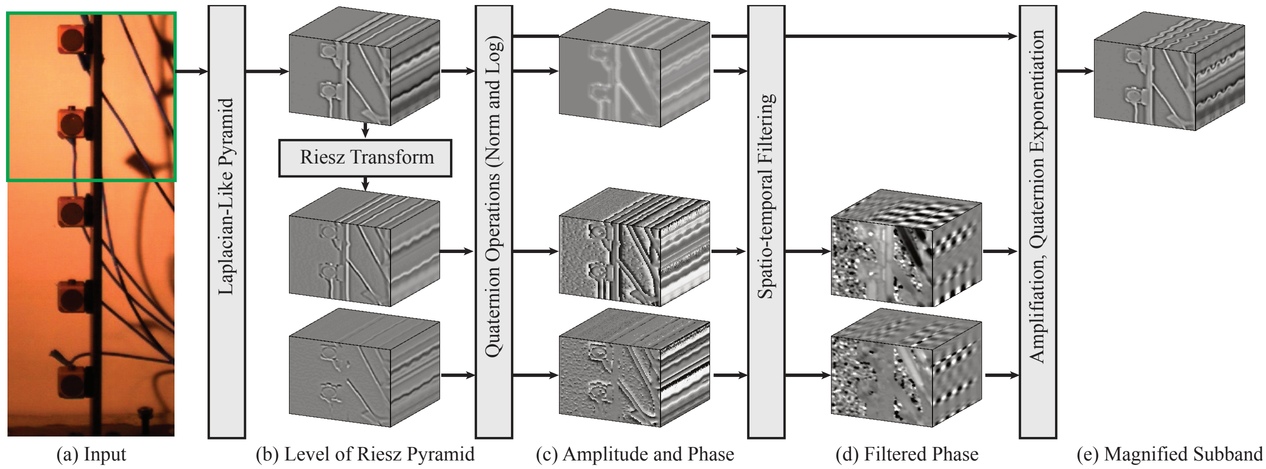


Figure 2. A processing pipeline for motion magnification with the Riesz pyramid. A region of interest (highlighted in green) of an input (a) is decomposed using a Laplacian-like pyramid (only one level shown). The Riesz transform of this level is taken to produce the Riesz pyramid (b). The quaternion norm is used to compute the amplitude (c, top row) and the quaternion logarithm is used to produce the quaternionic phase (c, bottom rows). The quaternionic phase is spatio-temporally filtered (d) to isolate motions of interest and then this quantity is used to phase-shift the input Riesz pyramid level to produce a motion magnified subband (e). These subbands can then be collapsed to produce a motion magnified video (not shown).

Spatial Smoothing We can perform a spatial amplitude weighted blur with Gaussian kernel K_ρ with standard deviation ρ on the i and j components of the temporally filtered signal to further increase SNR

$$i\frac{A\phi'\cos(\theta) * K_\rho}{A * K_\rho} + j\frac{A\phi'\sin(\theta) * K_\rho}{A * K_\rho} \tag{23}$$

where A is the amplitude of the Riesz pyramid coefficients. If we use our assumption that the orientation does not change substantially in the support of K_ρ, then we can move $\cos(\theta)$ and $\sin(\theta)$ outside of the convolution in Eq. 23 to get.

$$i\cos(\theta)\phi'' + j\sin(\theta)\phi'' \tag{24}$$

where $\phi'' = \frac{A\phi' * K_\rho}{A * K_\rho}$.

4.2. Amplification

We motion amplify a Riesz pyramid coefficient in the same way we would phase-shift a complex number. First, we perform a quaternion exponentiation on the filtered and amplified (by α) quaternionic phase (Eq. 24) to produce a unit quaternion

$$\cos(\alpha\phi'') + i\sin(\alpha\phi'')\cos(\theta) + j\sin(\alpha\phi'')\sin(\theta) \tag{25}$$

We then multiply this unit quaternion by the original coefficient $I + iR_1 + jR_2$ in the Riesz pyramid. We only need the real part of the result, which by Eq. 7 is equal to

$$I\cos(\alpha\phi'') - R_1\sin(\alpha\phi'')\cos(\theta) - R_2\sin(\alpha\phi'')\sin(\theta) \tag{26}$$

This gives the coefficients of a real Laplacian-like pyramid for every frame, in which the motions have been magnified, which can then be collapsed to produce a motion magnified video (Fig. 2(e)).

In Wadhwa et al. 2014 [7], we provided an alternate formulation of the magnification in terms of phase shifting the input subband when it is regarded as the real part of a complex number, whose imaginary part is a quadrature pair that is 90 degrees phase shifted along the dominant orientation θ. We now show that what we did in that paper produces the same results as what we present here.

As our first step in that work, we take Eq. 23 and combine the two components by multiplying by $\cos(\theta)$ and $\sin(\theta)$ and then summing to get

$$\cos(\theta)\frac{A\phi'\cos(\theta) * K_\rho}{A * K_\rho} + \sin(\theta)\frac{A\phi'\sin(\theta) * K_\rho}{A * K_\rho} = \phi'' \tag{27}$$

which is Eq. 15 in Wadhwa et al. 2014 [7]. We then amplify this quantity by α and then use complex exponentiation to get

$$\cos(\alpha\phi'') + i\sin(\alpha\phi''). \tag{28}$$

This is then multiplied by $I + iQ$ where $Q = R_1\cos(\theta) + R_2\sin(\theta)$ is a Hilbert transform of I along the dominant orientation θ. The multiplication yields

$$I\cos(\alpha\phi'') - R_1\sin(\alpha\phi'')\cos(\theta) - R_2\sin(\alpha\phi'')\sin(\theta) \tag{29}$$

which is identical to Eq. 26 showing that the quaternion formulation yields the same results as that specified in Wadhwa et al. 2014 [7].

5. Conclusion

We presented a quaternionic formulation of the Riesz pyramid. In this formulation, we use the quaternionic phase as a proxy for the motion signal. Because this representation is invariant to the sign ambiguity in the local phase of the signal, we can spatially and temporally filter it to isolate motions of interest. This can then be amplified and used to phase shift the original Riesz pyramid coefficients to produce a motion magnified signal.

References

[1] M. Felsberg and G. Sommer. The monogenic signal. *IEEE Trans. Signal Process.*, 49(12):3136–3144, 2001.

[2] J. Lee and S. Y. Shin. General construction of time-domain filters for orientation data. *Visualization and Computer Graphics, IEEE Transactions on*, 8(2):119–128, 2002.

[3] C. Liu, A. Torralba, W. T. Freeman, F. Durand, and E. H. Adelson. Motion magnification. *ACM Trans. Graph.*, 24:519–526, Jul 2005.

[4] M. Rubinstein. *Analysis and Visualization of Temporal Variations in Video*. PhD thesis, Massachusetts Institute of Technology, Feb 2014.

[5] M. Unser, D. Sage, and D. Van De Ville. Multiresolution monogenic signal analysis using the riesz–laplace wavelet transform. *IEEE Trans. Image Process.*, 18(11):2402–2418, 2009.

[6] N. Wadhwa, M. Rubinstein, F. Durand, and W. T. Freeman. Phase-based video motion processing. *ACM Trans. Graph.*, 32(4):80:1–80:10, July 2013.

[7] N. Wadhwa, M. Rubinstein, F. Durand, and W. T. Freeman. Riesz pyramid for fast phase-based video magnification. In *Computational Photography (ICCP), 2014 IEEE International Conference on*. IEEE, 2014.

[8] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. Freeman. Eulerian video magnification for revealing subtle changes in the world. *ACM Trans. Graph.*, 31(4):65:1–65:8, July 2012.