1. INTRODUCTION

Multiscale (MS) decomposition of images, e.g., by using wavelet transforms [1], [2], is known as an effective method for analyzing image signals. We can use it in various image processing applications, such as compression, denoising, enhancement, and texture retrieval. As a method for implementing MS decomposition, Laplacian pyramid (LP) [3] is widely used. The LP performs separable low-pass filtering along both horizontal and vertical directions, followed by the explicit downsampling.

Many images contain diagonal edges as well as horizontal/vertical ones. Therefore, multidirection (MD) decomposition is also beneficial as well as MS decomposition. Without being exhaustive, contourlet [4–7], directionlet [8,9], and curvelet [10] are key methods for MSMD decomposition based on the traditional MS decomposition. In particular, the contourlet transforms (CTs) are closely related to the LP, and achieve the MSMD decomposition by combining the LP and directional filter banks (DFBs) [11–13].

The retargeting pyramid (RP) [14, 15] has been proposed as an alternative to the LP. It achieves MS decomposition considering the importance of content in an image. Furthermore, it can be incorporated with DFBs to realize content-aware MSMD decomposition. The RP replaces low-pass filtering and downsampling in the LP with image retargeting, which is also known as content-aware image resizing [17–27]. Roughly speaking, the explicit low-pass filtering and downsampling approach in the LP is replaced with implicit downsampling and filtering by utilizing image retargeting. However, image retargeting in the RP requires a two-step interpolation: It leads to performance loss for image processing.

This paper proposes a novel decimation operator for the RP using the bilateral filter (BF) [28]. It replaces the two-step interpolation used in the RP with a single interpolation. In order to further improve the performance of the RP, we use Tikhonov regularization [29] to design the decimation matrix. The improved RP (iRP) using our proposed operator outperforms CTs with conventional pyramid structures in our experiments.

The remaining of this paper is constructed as follows. Section 2 gives a review of the RP. The iRP is introduced in Section 3. Section 4 presents some experimental results. Finally, Section 5 concludes this paper.

2. REVIEW

In this section, we briefly review the previous structure of the RP.

2.1. Retargeting Pyramid

Considering an $H_0 \times W_0$ image $X$, let $x^{(0)}(0) \in \mathbb{R}^{H_0 \times W_0}$ be the vectorized version of $X$. The significances (saliencies) of an image is transformed by retargeting. For the $K$-level decomposition, retargeting is performed by $K_{sig}$ levels, and its output is further transformed by using the LP for $K_{sig} = K - K_{sig}$ levels. The $k$-th level outputs $x^{(k+1)}$ and $\hat{x}^{(k)}$ in the RP are represented as

$$x^{(k+1)} = \begin{cases} \mathbf{R}^{(k)}x^{(k)}, & 0 \leq k \leq K_{sig} - 1 \smallskip \mathbf{L}^{(k)}x^{(k)}, & K_{sig} \leq k \leq K - 1 \end{cases}, \quad \hat{x}^{(k)} = \begin{cases} x^{(k)} - \mathbf{R}^{(k)}x^{(k+1)}, & 0 \leq k \leq K_{sig} - 1 \smallskip \mathbf{H}^{(k)}x^{(k)}, & K_{sig} \leq k \leq K - 1 \end{cases}$$

where $x^{(k+1)}$ indicates a low-pass component, $\hat{x}^{(k)}$ is a high-pass component, and $\mathbf{R}$ is the retargeting operator described in Sec. 2.2. $\mathbf{L}$ and $\mathbf{H}$ are the low-pass and high-pass filters, respectively, creating a perfect reconstruction filter bank. $\mathbf{R}$ contains low-pass filtering and downsampling implicitly, and $\mathbf{R}^{\dagger}$ is the pseudo inverse operator of $\mathbf{R}$. Note that $\mathbf{R}^{\dagger}\mathbf{R} \neq \mathbf{I}$ where $\mathbf{I}$ is an identity matrix. In the previous RP [14, 15], the structure is designed as shown in Fig. 1, and the retargeting operator $\mathbf{R}$ consists of two independent operations:

$$\mathbf{R} = \mathbf{A}\Phi,$$

where $\mathbf{A}$ is uniform bicubic scaling, and $\Phi$ is the warping operation which deforms original pixels while keeping the original size of $k$-th level input image.
The algorithm of image retargeting used in the RP is similar to a mesh warping method described in [20]. Let \( p \) be the original pixel position such as the warped image by \( \Phi \), thereby, is expected to be superior in edge preserving.

The flow of our proposed iRP is outlined in Fig. 2. It is constructed using a significance map. In the RP, the significance is represented as a weight in a similar way in [20]:

\[
W = \frac{1}{2} \left( \frac{1}{\max(\nabla X)} \right) \omega_X + W_h, \tag{5}
\]

where \( \omega_X = \| \nabla X \|_2 \) is the \( l_2 \)-norm of the gradient, and \( W_h \) is a saliency map which is given in [16]. In order to optimize a mesh based on (5), we consider the following cost function:

\[
E(p', s_i, l_i) = E_m(p') + \gamma \cdot (E_u(p', s_i) + E_l(p', l_i)), \tag{6}
\]

where \( \gamma \) is the weight for \( E_m \), \( s_i \) is the scale factor of a quad face between \( p \) and \( p' \), and \( l_i \) is the length ratio of the quad edges. In (6), \( E_m \) is used to stretch meshes based on the value of \( W \). \( E_u \) keeps the shapes of original mesh faces, and \( E_l \) prevents edges from being bent among vertices (pixels). The optimal mesh can be obtained by updating \( p' \) iteratively.

3. DIRECT DECIMATION FOR IMPROVED RP

A novel decimation operator for the RP is proposed in this section. The flow of our proposed iRP is outlined in Fig. 2. It is constructed by replacing the two-step interpolation in the previous RP with a single interpolation that can be obtained by modifying the BF. The iRP, thereby, is expected to be superior in edge preserving.

The filter coefficients of the modified BF are represented pixels to be interpolated and gray ones are available pixels for interpolation.
as follows:

\[ B^{(k)}(j, v_{f,j}(n)) = \frac{w_{f,i}(j,n)w_{f,j}(n)}{\sum_{n=mn} w_{f,i}(j,n)w_{f,j}(n)}, \]  

(11)

\[ w_{f,i}(j,n) = \exp \left( -\frac{d_{f,x}(j,n)^2 + d_{f,y}(j,n)^2}{2\sigma_1^2} \right), \]

(12)

\[ w_{f,j}(j,n) = \exp \left( -\frac{(r_f(j) - x^{(k)})(v_{f,j}(n))^2}{2\sigma_2^2} \right), \]

(13)

\[ d_{f,x}(j,n) = q^{(k)}_{d,x}(v_{f,j}(n)) - q^{(k)}_{v,x}(j)/s, \]

(14)

\[ d_{f,y}(j,n) = q^{(k)}_{d,y}(v_{f,j}(n)) - q^{(k)}_{v,y}(j)/s, \]

(15)

where \( j \) is a pixel index of \( x^{(k+1)} \) in (9) and \( v_{f,j} \) contains pixel indices in the support region of the BF, which satisfies both (16) and (17):

\[ q^{(k)}_{u,x}(j)/s - l_f \leq q^{(k)}_{d,x}(v_{f,j}(n)) \leq q^{(k)}_{u,x}(j)/s + l_f, \]

(16)

\[ q^{(k)}_{u,y}(j)/s - l_f \leq q^{(k)}_{d,y}(v_{f,j}(n)) \leq q^{(k)}_{u,y}(j)/s + l_f, \]

(17)

in which \( l_f \) is the filter size of \( B^{(k)} \) and \( n \) is the index of \( v_{f,j} \) illustrated in Fig. 3. Additionally, \( n_{mn} \) in (11) is the number of elements in \( v_{f,j} \). \( q^{(k)}_{d,x} \) and \( q^{(k)}_{d,y} \) are vectors containing coordinates of \( p' \) in (4), respectively, and \( \sigma_1 \) and \( \sigma_2 \) are standard deviations of Gaussian distributions.

\( B^{(k)} \in \mathbb{R}^{H_kW_k \times H_kW_k} \) in (8) can be similarly obtained as follows. First, a blurred image \( r_b \) is calculated as

\[ r_b = G^{(k)}x^{(k)}, \]

(18)

where \( G^{(k)} \) is the Gaussian filtering operator. After that, \( B^{(k)} \) is obtained as

\[ B^{(k)}(i, v_{b,i}(m)) = \sum_{m=mm} w_{b,i}(i,m)w_{b,j}(i,m), \]

(19)

\[ w_{b,i}(i,m) = \exp \left( -\frac{d_{b,x}(i,m)^2 + d_{b,y}(i,m)^2}{2\sigma_1^2} \right), \]

(20)

\[ w_{b,j}(i,m) = \exp \left( -\frac{(r_b(i) - x^{(k+1)}v_{b,j}(m))^2}{2\sigma_2^2} \right), \]

(21)

\[ d_{b,x}(i,m) = q^{(k)}_{d,x}(i) - q^{(k)}_{v,x}(v_{b,i}(m))/s, \]

(22)

\[ d_{b,y}(i,m) = q^{(k)}_{d,y}(i) - q^{(k)}_{v,y}(v_{b,i}(m))/s, \]

(23)

where \( i \) is shown in (4) and \( v_{b,i} \) contains pixel indices in the support region of the BF. It satisfies both (24) and (25):

\[ q^{(k)}_{d,x}(i) - l_b \leq q^{(k)}_{u,x}(v_{b,i}(m))/s \leq q^{(k)}_{d,x}(i) + l_b, \]

(24)

\[ q^{(k)}_{d,y}(i) - l_b \leq q^{(k)}_{u,y}(v_{b,i}(m))/s \leq q^{(k)}_{d,y}(i) + l_b, \]

(25)

where \( l_b \) is the filter size of \( B^{(k)} \) and \( m \) is the pixel index of \( v_{b,i} \). \( v_{b,i}(m) \) is also illustrated in Fig. 3. Furthermore, \( n_{mm} \) in (19) is the number of elements in \( v_{b,i} \).

3.3. Derivation of \((B^{(k)})^+\) using Tikhonov Regularization

Tikhonov regularization [29] is further applied to yield a decision matrix. \((B^{(k)})^+\) in (7) can be obtained by solving the following least squares problem:

\[ \|B^{(k)}x^{(k+1)} - x^{(k)}\|^2 + \lambda\|x^{(k+1)}\|^2, \]

(26)

where \( \lambda \) is an arbitrary weight. As a result, \((B^{(k)})^+\) is obtained as

\[ (B^{(k)})^+ = (B^{(k)})^T(B^{(k)})^T + \lambda I)^{-1}(B^{(k)})^T. \]

(27)

4. EXPERIMENTAL RESULTS

In this section, we validate our proposed iRP by applying it to a couple of image processing applications. We used three 256 × 256
grayscale images of Monarch, Pepper, and Lena as the test images. As the conventional methods, we use the discrete wavelet transform (WT), bicubic resizing (BC) and pyramid in [4] (MD). For the WT and the low-pass filter for the LP, the well-known biorthogonal 9/7 (WT), bicubic resizing (BC) and pyramid in [4] (MD). For the WT

As the conventional methods, we use the discrete wavelet transform (WT), bicubic resizing (BC) and pyramid in [4] (MD). For the WT

4.2. Image Denoising

Here, we introduce the performance of the CT based on the iRP (referred to CT-iRP) by applying it to image denoising. It can be obtained by replacing a pyramid in the CT with the iRP and directional subbands are calculated as

\[ \tilde{y}^{(k)} = D_n^{(k)} \hat{x}^{(k)}, \]  

(28)

where \( \tilde{y}^{(k)} \) and \( D_n^{(k)} \) are directional subbands and the \( k \)-th level transform by DFBs [11–13] with \( 2^n \) subbands, respectively. For comparing the pure performance of pyramid structures, we choose the simple hard-thresholding method. We select a threshold of \( 3\sigma \), where \( \sigma \) is the standard deviation of noise. Furthermore, redundancies of all CTs are set to about 1.33. There were \{32, 16, 16, 0, 0\} directional subbands for the CT with the LP [5] (referred to as CT-LP), \{32, 16, 8, 4, 4\} for the CT with the improved pyramid structure [4] (referred to as CT-MD), and \{8, 16, 8, 4, 4\} for the CT-RP [14, 15] and the CT-iRP from fine to coarse scale. The numbers for the CT-LP and the CT-MD were set as those reported by the original researchers. The same filters as those in the original research on the CT [4, 5], e.g., PKV A 2-D filter bank [13], were used. Table 3 summarizes the performance of denoising. It shows that the CT-iRP always outperforms the CT-LP and all methods within \( 5 \leq \sigma \leq 30 \) for all test images. The enlarged portions of denoised images of Monarch are shown in Fig. 5. It shows distortions that cover up the whole images due to the DFB decomposition are reduced by using the CT-iRP. In addition, the CT-iRP preserves edges as shown in white dots in wings.

5. CONCLUSION

In this paper, we propose a novel decimation operator for the retargeting pyramid. We replace the two-step interpolation in the previous RP with a single interpolation using the customized bilateral filter. Furthermore, the decimation operator is improved by using Tikhonov regularization. In our experimental results, the RP using the proposed method outperformed conventional MSMD decompositions for a few image processing applications. Our future work includes reducing the computational complexity and investigating a method to compress side information of mesh in order to apply the iRP to image coding.

6. REFERENCES


