Pan-sharpening of multi-spectral images using over-complete rational-dilation wavelet transform

Haijiang Wang (王海江)[1], Qinke Yang (杨勤科)[2*], Chunmei Wang (王春梅)[1], and Weiling Guo (郭伟玲)[1]

[1] *Institute of Soil and Water Conservation, Chinese Academy of Sciences and Ministry of Water Resources, Xianyang 712100, China*

[2] *Department of Urban and Resource Sciences, Northwest University, Xi'an 710069, China*

[*] *Corresponding author: qkyang@nwu.edu.cn*

Received December 16, 2011; accepted January 6, 2012; posted online May 9, 2012

A rational-dilation wavelet transform (RWT) based pan-sharpening method for multi-spectral (MS) images of various oscillatory nature is proposed. The previous multi-scale transforms, such as wavelets, curvelets and contourlets, decompose an image into channels with low constant Q-factors, and aren't suitable for pan-sharpening images with different behavior in frequency domain. The RWT as an over-complete scheme not only increases the sampling in spatial and frequency domain, but also provides a tunable Q-factor approach to be suitable for a given dataset. We studied its multi-scale decomposition scheme and the RWT based pan-sharpening method. The MS image pan-sharpening experiments show that this method using a better suitable parameter set can achieve a promising performance and often outperforms many other widely-used pan-sharpening methods both in visual quality and in term of evaluation indexes.

OCIS codes: 100.0100, 100.2000, 100.7410.

doi: 10.3788/COL201210.S11008.

Pan-sharpening is known as the injection of fine spatial information from the high spatial resolution panchromatic (PAN) image into the low spatial resolution multi-spectral (MS) images to get high spatial resolution MS images. This technique is of special interest in remote sensing, because satellites usually take PAN and MS images separately and high spatial and spectral resolution are both necessary for various complex tasks in land-cover classifications. Many studies proposed that multi-scale transforms (MST) are very useful for pan-sharpening or fusing MS image, such as wavelet transform (WT)[1−4] and non-separable WT[5,6]. Also, curvelet transform (CVT)[7−9] and contourlet transform[10,11] (CT), as two more efficient multi-scale and directional representation approach over the WT, were latterly introduced to the area and showed more desirable performances. Moreover, some studies show that the non-subsampled counterparts of some of these MSTs can further improve pan-sharpening quality, such as non-sampled WT[12] (NSWT) and non-sampled CT[13,14] (NSCT). However, these previous MST based methods should be improved for MS images of different oscillatory nature.

The low-Q multi-scale decomposition scheme (the multi-scale decomposition scheme with a low Q-factor) of the previous MST methods above is not suitable for various images with different behavior in frequency domain. In the radial direction of frequency domain, the previous MST methods iteratively decompose a signal into channels having a same low Q-factor, and belong to constant low-Q transforms. Figure 1 illustrates this decomposition scheme by taking the WT as an example. That is, the bandwidth of the band-pass filter with a little high central frequency will be too large to lead to a sparse distribution of these filters in high frequency area. This decomposition scheme or the Q-factor is only suitable for analyzing little oscillatory signal, but not for relatively more oscillatory signal[15,16]. MS images often contain more/less oscillatory features that may cause rich high/low frequency component in frequency domain, thus require a tunable-Q transform alternative.

Some approaches have been proposed to obtain the tunable-Q scheme. For example, Oppenheim and Makur *et al.* provided a mapping method that maps a discrete sequence to another specific sequence to achieve an unequally spaced frequency sampling[17,18]. Diniz *et al.* advised a kind of "bounded-Q fast filter bank" that splits the spectrum to obtain a piecewise linear frequency decomposition[19]. However, these methods neither provide a really tunable-Q factor, nor offer an efficient inverse transform. Recently, Bayram proposed a kind of over-complete rational-dilation wavelet transform (RWT) that achieves a constant-Q analysis with a really-tunable Q-factor. Compared to that of the previous MSTs, its Q-factor can be made lower/higher and the frequency resolution can be made sparser/finer[15,16]. Also, it is a tight frame and its inverse can be implemented efficiently via the transpose of the forward

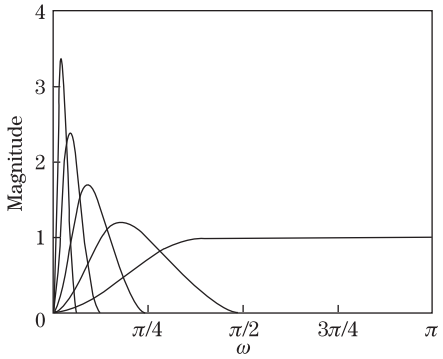


Fig. 1. Frequency-domain decomposition of WT.

transform. Subsequent studies have shown this MST method as a more desirable signal representation approach over the WT in some application area[16,20]. In this letter, we firstly gave a brief introduction to this RWT, and then presented the RWT based pan-sharpening algorithm. We also showed the desirably practical performance of this pan-sharpening method with experiments, and summarized the shortcomings and subsequent tasks of this method finally.

Most previous research on wavelet transforms with rational dilation considered only the critically-sampled case, such as the well-known orthogonal case of Auschor[15]. The critically-sampled RWT increases the sampling in frequency domain over the WT, but not in spatial domain. As shown in Fig. 2, the WT and NSWT can be understood to sample the time-frequency (T-F) plane with $\{A=2,\ B=2\}$ and $\{A=2,\ B=1\}$ respectively, whereas the critically-sampled RWT sets $\{A = q/p,\ B = q/(q-p)\}$. In contrast, the over-complete RWT, realized using the iterated filter bank (FB) in Fig. 3, samples the T-F plane more densely with $\{A = q/p,\ B = s\}$. Here and below, $p,\ q \in$Z, $1 \leqslant p < q$, gcd$(p,\ q)=1$ and $s \in$Z, $p/q+1/s \geqslant 1$ (the reason is presented below). Using the over-complete RWT, as seen in Fig. 3 and below, one can flexibly adjust the T-F sampling by suitably tuning the parameter set $(p,\ q,\ s)$.

We denote discrete-time sequences by lower case letters as $f(n)$ with $n \in$Z, and denote the discrete-time Fourier transform of $f(n)$ as $F(\omega)$. $F(\omega)$ is 2π periodic in ω, and is given by $F(\omega) = \sum_n f(n)\exp(-\mathrm{i}\omega n)$. Supposed all discrete-time signals are real valued, we have $F^*(\omega) = F(-\omega)$. The perfect reconstruction (PR) of the FB in Fig. 3 can be granted if

$$\begin{cases} H(\omega) = 0, & \omega \in [\pi/q,\ \pi] \\ G(\omega) = 0, & \omega \in [0,\ (1-1/s)\pi] \end{cases}, \text{ and}$$

$$\begin{cases} |H(\omega)| = \sqrt{pq}, & \omega \in [0,\ (1-1/s)\pi/p] \\ |G(\omega)| = \sqrt{s}, & \omega \in [p/q\pi,\ \pi] \end{cases}. \quad (1)$$

The transition bands of $H(\omega)$ and $G(\omega)$ can be arbitrarily specified respectively. However, both should keep a certain degree of differentiability so as to provide $h(n)$ and $g(n)$ with better time-domain localization ability at the same time. As higher differentiability in frequency domain results in the slower decay of the impulse response in time domain, the frequency response of Daubechies wavelet with the vanishing moment number of two was used here. As shown in Fig. 4, such specification provides the filter with a better localization ability in both time and frequency domain. After the specification of $H(\omega)$ and $G(\omega)$, the filtering to any discrete signal as that in Fig. 3 can be directly implemented in frequency domain.

In order to clarify the behavior of the iterative filter bank in Fig. 3, the order of the filters and the re-samplers was exchanged by using noble identities for obtaining the equivalent combined filter at each decomposition level that acts directly on the input and follows by a re-sampler (the band-limiting condition (1) of the designed low-pass filter makes the exchange possible)[16]. Finally, we obtained a low-pass combined filter $\overline{\Phi}_{\mathrm{j}}(\omega)$ and a high-pass

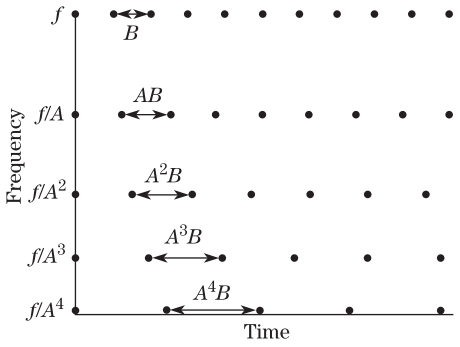


Fig. 2. Sampling lattice in T-F plane.

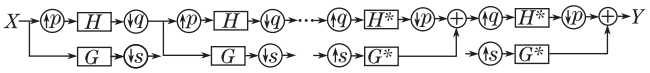


Fig. 3. Decomposition and reconstruction of RWT.

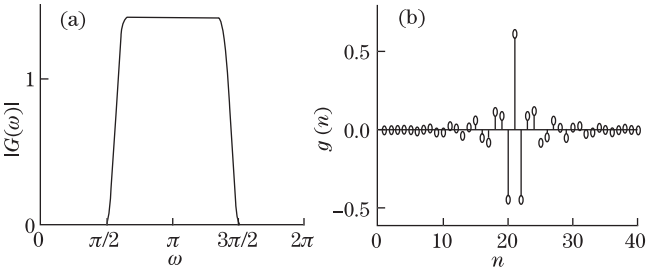


Fig. 4. $|G(\omega)|$ (a) and $g(n)$ (b) in the case of $(p,\ q,\ s)=(2,\ 3,\ 2)$.

combined filter $\overline{\Psi}_{\mathrm{j}}(\omega)$ for the decomposition level of j. For $j=1$, $\overline{\Phi}_1 = H(\omega)/p$ and $\overline{\Psi}_1(\omega) = G(\omega)$; for $j \geqslant 2$, $\overline{\Phi}_{\mathrm{j}}(\omega) = \Phi_{\mathrm{j}}(\omega/p^j)/p^j$, and $\overline{\Psi}_{\mathrm{j}}(\omega) = \Psi_j(\omega/p^{j-1})/p^{j-1}$, $\omega \in [0,\ \pi]$. Here,

$$\Phi_{\mathrm{j}}(\omega) = \begin{cases} \prod_{k=0}^{j-1} H(\omega q^{j-(k+1)}p^k), & \omega \in [0,\ \pi/q^j] \\ 0, & \omega \in (\pi/q^j,\ \pi] \end{cases}, \text{ and}$$

$$\Phi_{\mathrm{j}}(\omega) = \begin{cases} \prod_{k=0}^{j-1} H(\omega q^{j-(k+1)}p^k), & \omega \in [0,\ \pi/q^j] \\ 0, & \omega \in (\pi/q^j,\ \pi] \end{cases}. \quad (2)$$

Using the two combined filters, one can convert the iterative filter bank in Fig. 3 to its equivalent filter bank as presented in Fig. 5. Here, each structure surrounded by an ellipse is an ideal re-sampler[16]. We then can calculate the Q-factor as a function of the parameter set $(p,\ q,\ s)$. From the analysis above, we see both $\overline{\Phi}_{\mathrm{j}}(\omega)$ and $\overline{\Psi}_{\mathrm{j}}(\omega)$ has a flat pass band if $(1-1/s) > (p/q)^2$. Provided that the two transition bands of the $\overline{\Psi}_{\mathrm{j}}(\omega)$ are both half-band, just like the specification here, the edge of the pass band on each side is the middle point of each transition band respectively. Then we can obtain that $Q_j(p,\ q)=\sqrt{p/q}/(1-p/q)$ for $j > 1$.

Since $Q_j(p,\ q)$ is constant and independent of the decomposition level j, the iterated FB shown in Fig. 3 is a constant-Q transform. As shown in Fig. 6, one can flexibly achieve a sparser/finer frequency-domain

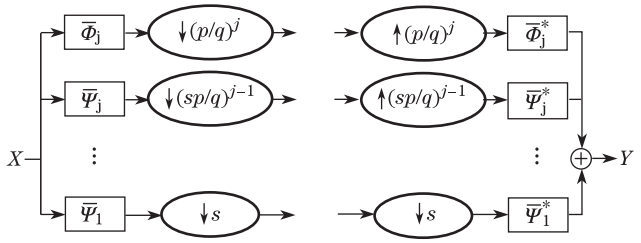


Fig. 5. Equivalent system of that in Fig. 3.

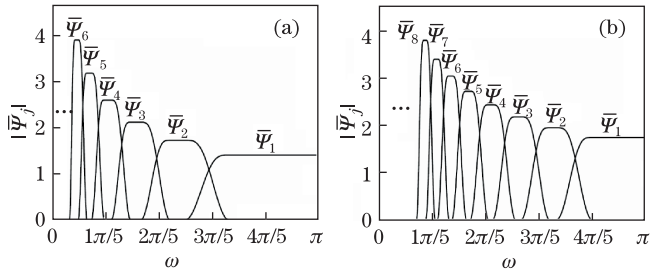


Fig. 6. $|\overline{\Psi}_j|$ of RWT in different (p, q, s) sets: (a) (2, 3, 2) and (b) (4, 5, 3).

decomposition by using a low/high Q-factor. Because sparser/finer frequency-domain decomposition results in finer/sparser space-domain decomposition, one should rationally pursue the Q-factor in practical.

The one-dimensional (1D) over-complete RWT discussed above can be expanded to its 2D counterpart via tensor product approach as other MST does[15]. As seen in Figs. 7 and 2, we should suitably choose the parameter s, since there is an inverse relationship between the ideally frequency-domain response and the densely spatial-domain sampling.

The RWT (simplified denotation for the over-complete RWT) based pan-sharpening algorithm is similar to that of many other MSTs[1−4,10,11]. Step I, register the MS image and PAN image; perform principle-component-analysis (PCA) or intensity-hue-saturation (IHS) transform to the original MS image; match the histogram of PAN image (P) to the first component or the intensity (I) of the MS image; obtain the matching result (P'). Step II, fuse I and P' to generate fused image I'; perform J-levels RWT decomposition to P' and I as that in Fig. 3; obtain one approximation and J details for P' and I respectively; fuse the two approximations based on a kind of fusion rule, and each detail pair based on the same rule or another; perform RWT reconstruction to the fused approximation and details to obtain I'. Step III, perform PCA or IHS inverse transform using I' to generate the pan-sharpened MS image.

We pan-sharpened the MS images of ETM+ datasets here to compare this RWT based pan-sharpening method with other methods and evaluate the effectiveness of the tunable-Q factor in the RWT. Two ETM+ datasets of different oscillatory nature in Yantai, China, were used here. As shown in Fig. 8, each MS image is composed of Band2, Band3 and Band4 and has a size of 256×256×3, and each PAN image has a size of 512×512.

Four other MST based pan-sharpening methods were also performed for comparison, including that of WT, NSWT, CVT and NSCT. In addition, IHS and PCA as two traditional methods were considered too. We employed four-level decomposition for all the MSTs, and used Daubechies 9/7 filters for the WT and NSWT. For the directional FB in the CVT and NSCT, we used (3, 4, 4, 5) directional levels for the four levels respectively. Different parameter sets were taken into account for the RWT, whereas the conditions $q = p + 1$ and $s = p$ were also used for a clear comparison. Therefore, we used RWT(p) to denote these RWTs using different parameter sets below.

For all these MST based pan-sharpening methods, we used the IHS transform to obtain the component I, and used the same set of fusion rules as described below. These fusion rules are relatively simple, as this way, one can observe to what extent the method is efficient without the use of more complex fusion rules. The approximation of the fused image I' was replaced by the average of the two approximations obtained from P' and I respectively, and the detail of I' at each level was obtained by using the following three fusion rules respectively. Rule I is based on pixel value, i.e. to set the corresponding larger absolute value of the two details (resulted from P' and I respectively) to be the detail of I'. Rule II is based on regional absolute value, i.e. firstly to compare the regional absolute value of the two details and obtain a binary decision map, then to perform consistency verification to the map, finally to obtain the detail of I' based on the verified map[1,3]. Rule III is based on regional variance. It is the same as Rule II except using the regional variance rather than the regional absolute value. In this experiment, the region for the Rule II and III is a 3×3 window for each pixel.

We also used several fusion quality indexes to objectively evaluate the performance of each method, including the standard variance[3,14] (Std), the entropy[10,14] (E) and the sharpness[10,14] (SP) of the pan-sharpened images, as well as the spectral distortion[4] (Dist), the correlation[5,12] (Corr), the relative average spectral error[2,8] (RASE) and the relative global dimensional synthesis error[2,8,9] (ERGAS) between original images and pan-sharpened images. The former three of the seven indexes will be higher if more structural detail is preserved; the Dist, RASE and ERGAS will be lower and

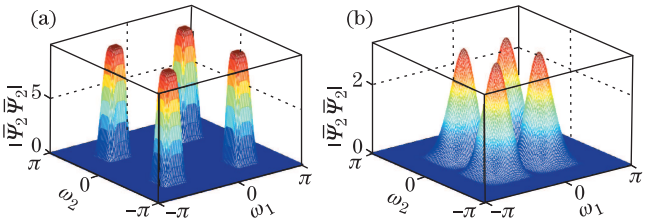
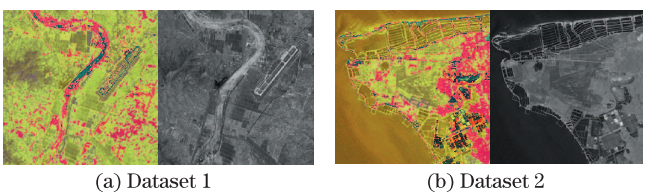


Fig. 7. $|\overline{\Psi}_2\overline{\Psi}_2|$ of 2D RWT in different (p, q, s) sets: (a) (2, 3, 2) and (b) (2, 3, 1).



(a) Dataset 1 (b) Dataset 2

Fig. 8. (a) Composed MS images of Band 2, Band 3 and Band 4, and (b) the PAN images, for the two datasets used.

the Corr will be higher if less spectral distortion is generated. We implemented all the experiments in MATLAB environment.

For both datasets, as shown in Figs. 9 and 10, PCA and IHS generate severer spectral distortion compared with all the MST based methods, and RWT based method preserves slightly more structural details compared with the other MST based methods (the reason to use such two parameters is presented below). Moreover, the pan-sharpened results based on the other two rules show similar advantages of the RWT based method. Using the proposed method, we can see the spatial and the spectral resolution of the original MS images appear to have been better enhanced. That is, the results of the pan-sharpening contain more structural details of the PAN image and richer spectral information of the MS images.

Tables 1 and 2 present a detailed and objectively comparison among these methods. For Dataset 1 and all the fusion rules, the RWT using the parameter 3 obtains more desirably quality index results than the others in most cases, including not only the two traditional methods and the other MST based methods, but also these using the other parameters; for Dataset 2, the one using the parameter 5 achieves the similar performance.

To evaluate the effectiveness of the tunable Q-factor in the proposed method, we presented the values resulted from the RWT using different parameters and Rule III in Fig. 11. As seen, for a give dataset, there exist better suitable (p, q, s) parameter sets that can yield more promising pan-sharpened result among all the parameter set cases. Furthermore, the values corresponding to the other two rules also share this variation pattern.

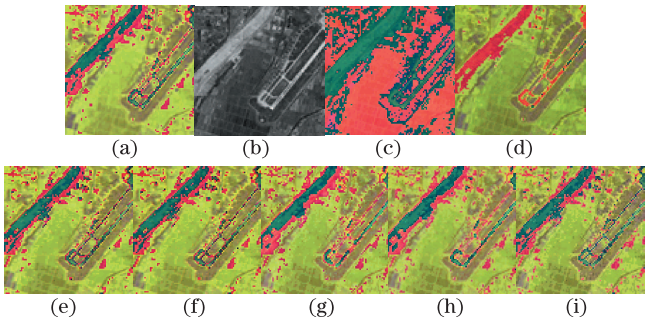


Fig. 9. Close-up show of the original (a) MS and (b) PAN images, and the results from the method based on (c) IHS, (d) PCA, (e) WT, (f) NSWT, (g) CVT, (h) NSCT, and (i) RWT(3), using Rule III and Dataset 1.

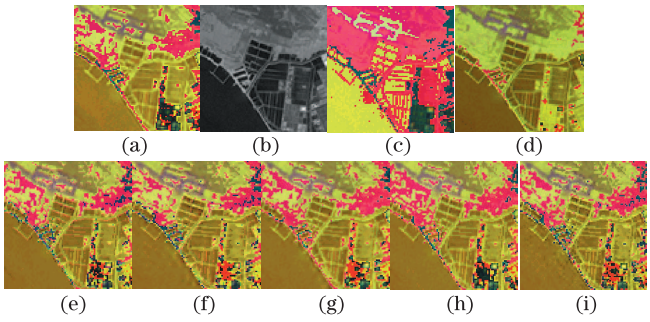


Fig. 10. Close-up show of the original (a) MS and (b) PAN images, and the results from the method based on (c) IHS, (d) PCA, (e) WT, (f) NSWT, (g) CVT, (h) NSCT, and (i) RWT(5), using Rule III and Dataset 2.

Table 1. Fusion Quality Indexes from Different Methods for Dataset 1

	Method	Std	Ent	SP	Dist	Corr	RASE (%)	ERGAS
	WT	19.713	6.977	10.718	5.455	0.933	4.415	1.981
	NSWT	19.719	6.978	10.721	5.452	0.932	4.374	1.852
Rule	CVT	19.738	6.992	10.752	5.446	0.934	4.133	1.780
I	NSCT	19.723	6.985	10.756	5.447	0.934	4.087	1.742
	RWT(3)	19.741	7.004	10.750	5.441	0.935	3.968	1.684
	RWT(5)	19.708	6.975	10.715	5.486	0.930	4.484	1.992
	WT	19.358	6.966	10.539	5.536	0.924	4.382	1.909
	NSWT	19.363	6.968	10.541	5.538	0.924	4.119	1.822
Rule	CVT	19.377	6.974	10.550	5.535	0.925	3.935	1.685
II	NSCT	19.381	6.976	10.552	5.535	0.925	3.981	1.703
	RWT(3)	19.379	6.977	10.559	5.521	0.925	3.894	1.634
	RWT(5)	19.349	6.964	10.536	5.566	0.923	4.356	1.886
	WT	19.378	6.972	10.553	5.640	0.925	4.453	1.954
	NSWT	19.380	6.974	10.558	5.610	0.925	4.272	1.847
Rule	CVT	19.402	6.978	10.559	5.594	0.926	3.933	1.716
III	NSCT	19.407	6.980	10.560	5.592	0.926	3.928	1.704
	RWT(3)	19.412	6.983	10.562	5.566	0.927	3.920	1.656
	RWT(5)	19.373	6.974	10.559	5.595	0.926	4.364	1.894
	IHS	13.714	6.874	6.622	18.341	0.908	26.566	10.502
	PCA	7.988	6.803	4.427	36.762	0.871	49.752	21.629

The values in the bold-font style are the best results for each fusion rule.

Table 2. Fusion Quality Indexes from Different Methods for Dataset 2

	Method	Std	Ent	SP	Dist	Corr	RASE (%)	ERGAS
	WT	31.856	6.825	12.492	7.724	0.926	6.491	2.846
	NSWT	31.975	6.832	12.504	7.719	0.928	6.469	2.814
Rule	CVT	32.057	7.023	12.516	7.638	0.928	6.225	2.732
I	NSCT	32.024	6.849	12.531	7.646	0.928	6.162	2.691
	RWT(3)	31.618	6.821	12.487	7.794	0.927	6.574	2.933
	RWT(5)	32.159	7.019	12.527	7.622	0.929	5.911	2.471
	WT	31.164	6.723	11.582	8.017	0.915	6.016	2.471
	NSWT	31.276	6.725	11.586	8.023	0.915	5.964	2.425
Rule	CVT	31.372	6.894	11.589	7.912	0.916	5.526	2.288
II	NSCT	31.39	6.826	11.593	7.984	0.916	5.541	2.357
	RWT(3)	31.043	6.718	11.579	8.118	0.914	6.114	2.482
	RWT(5)	31.413	6.915	11.608	7.913	0.916	5.531	2.309
	WT	31.486	6.714	11.328	8.122	0.913	6.117	2.604
	NSWT	31.54	6.728	11.369	8.110	0.914	6.024	2.581
Rule	CVT	31.552	6.849	11.488	8.003	0.917	5.804	2.424
III	NSCT	31.554	6.832	11.505	7.996	0.916	5.765	2.398
	RWT(3)	31.101	6.805	11.356	8.135	0.915	6.281	2.633
	RWT(5)	31.577	6.942	11.612	7.936	0.917	5.528	2.307
	IHS	25.774	6.818	7.941	22.744	0.910	38.410	16.497
	PCA	15.958	6.634	5.224	32.651	0.884	51.954	24.583

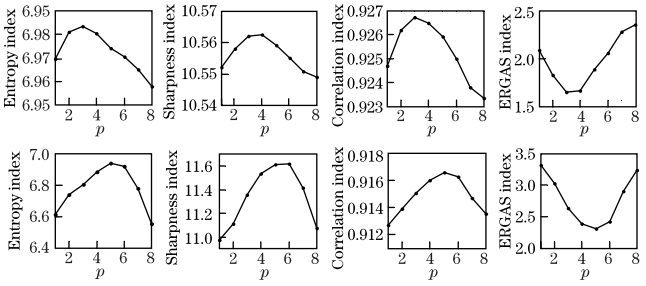


Fig. 11. Quality indexes resulted from the RWT using different parameters, for Dataset 1 (upper) and Dataset 2 (down).

In Conclusion, we propose a rational-dilation wavelet based pan-sharpening scheme for pan-sharpening MS images of different oscillatory nature. The transform provides us with a tunable parameter set of (p, q, s) by which one can flexibly tune the Q-factor in the multiscale decomposition. We employed the RWT based pan-sharpening method to fuse the MS images and PAN images of ETM+ datasets. The results show that for this pan-sharpening method, different (p, q, s) sets will result in different performance, and these using better suitable (p, q, s) sets preserve more feature details while generate less spectral distortion compared with that of other widely-used pan-sharpening methods.

As seen, the curvelet and contourlet based pan-sharpening methods sometimes outperformed the proposed method, so an effectively directional-selective scheme should be integrated into this method latterly. Moreover, one should take into account the best suitable parameter set for a given dataset when using the method in practice, thus another subsequent issue is to find the strategy of choosing the best suitable (p, q, s) set corresponding to a particular dataset.

This work was supported by the National Natural Science Foundation of China (No. 2007CB407203).

References

1. H. Li, B. S. Manjunath, and S. K. Mitra, Graphical Models and Image Processing **57,** 235 (1994).
2. M. Gonzalez-Audicana, J. L. Saleta, R. G. Catalan, and R. Garcia, IEEE Trans. Geoscience and Remote Sensing **42,** 1291 (2004).
3. G. Pajares and J. M. D. L. Cruz, Pattern Recognition **37,** 1855 (2004).
4. J. Zhou, D. L. Civco, and J. A. Silander, International J. Remote Sensing **19,** 743 (2010).
5. H. Wang, Z. Jing, and J. Li, Chin. Opt. Lett. **1,** 523 (2003).
6. B. Liu, W. Liu, and J. Peng, Chin. Opt. Lett. **8,** 384 (2010).
7. E. Candeès, L. Demanet, D. Donoho, and L. Ying, Multiscale Modeling Simulation **5,** 861 (2006).
8. M. Choi, R. Y. Kim, M. R. Nam, and H. O. Kim, IEEE Geosci. Remote Sensing Lett. **2,** 136 (2005).
9. F. Nencini, A. Garzelli, S. Baronti, and L. Alparone, Information Fusion **8,** 143 (2007).
10. Q. Miao and B. Wang, in *Proceedings of International Society for Optical Engineering* **4,** 624 20Z.1 (2006).
11. Y. Zheng, C. Zhu, J. Song, and X. Zhao, in *Proceedings of IEEE International Conference on Information Acquisition* 420 (2006).
12. J. Nunez, X. Otazu, O. Fors, A. Prades, V. Pala, and R. Arbiol, IEEE Trans. Geoscience and Remote Sensing **37,** 1204 (1999).
13. Q. Zhang and B. L. Guo, Signal Processing **89,** 1334 (2009).
14. W. Kong, Y. Lei, Y. Lei, and X. Ni, IET. Signal Process **5,** 75 (2011).
15. I. Daubechies, *Ten Lectures on Wavelets* (SIAM, Philadelphia, 1992).
16. I. W. Selesnick, IEEE Trans. Signal Processing **99,** 1 (2011).
17. A. V. Oppenheim, D. H. Johnson, and K. Steiglitz, in *Proceedings of IEEE* **59,** 299 (1971).
18. A. Makur and S. K. Mitra, IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications **48,** 1086 (2001).
19. F. C. C. B. Diniz, I. Koethe, S. L. Netto, and L. W. P. Biscainho, EURASIP J. Advances in Signal Processing 94704 (2007).
20. I. W. Selesnick, in *Proceedings of the International Society for Optics and Photonics* 8138 (2011).