A Directional and Shift-Invariant Transform Based on M-channel Rational-Valued Cosine-Sine Modulated Filter Banks

Seisuke Kyochi*, Taizo Suzuki† and Yuichi Tanaka‡
*The University of Kitakyushu, Japan, Email: s-kyochi@kitakyu-u.ac.jp
† Nihon University, Japan, Email: taizo@ee.ce.nihon-u.ac.jp
‡ Tokyo University of Agriculture and Technology, Japan, Email: ytnk@cc.tuat.ac.jp

***Abstract*—This paper proposes a directional and shift-invariant transform based on M-channel rational-valued cosine-sine modulated filter banks (R-CSMFBs) for the practical implementation on hardware devices. M-channel CSMFBs can be easily designed by the modulation of a prototype filter and achieve a good stopband attenuation. In addition, in our previous work, the directionality and the shift-invariance of CSMFBs have been theoretically clarified. Thus, they can be an alternative choice of the dual-tree complex wavelet transform (DTCWT) which is one of the most popular directional and shift-invariant transforms. In this paper, it is shown that the proposed lifting-based structure of the R-CSMFB can also achieve rich directional selectivity and the shift-invariance even if the lifting coefficients are rounded to rational values. Finally, the R-CSMFB can provide better stopband attenuation and image denoising performance than that of the conventional M-channel rational-valued DTCWT in the simulation.**

I. Introduction

The dual-tree complex wavelet transform (DTCWT), conventionally, have been introduced to achieve "rich directional selectivity" and "shift-invariance" which discrete wavelet transforms (DWTs) cannot satisfy [1]. It can achieve better performance than DWTs in various kinds of image processing, such as image denoising [1].

In this paper, we address on the design of the directional and shift-invariant complex WT (CWT) satisfying the number of channel $M > 2$ and the rational value constraint simultaneously. Conventionally, an efficient design method for 2-channel rational-valued DTCWT (R-DTCWT) has been proposed for the hardware implementation (add/bit-shift operations) [2]. However, it has not been proposed yet for the M-channel ones. Although M-channel R-DTCWT can be realized by using the dual-tree complex wavelet packet (DTCWP) [3], a cascade of the 2-channel R-DTCWT suffers from poor stopband attenuation owing to the small number of the design parameters, especially in the case of the short filter length constraints on the entire system.

In our previous work, the directional selectivity and the shift-invariance of cosine-sine modulated filter banks (CSMFB) [4] have been theoretically clarified [5]. Thus, it can be considered that CSMFBs are in the same class of the directional and shift-invariant CWTs where the DTCWT belongs to. In this paper, the rational-valued CSMFB (R-CSMFB) is proposed by introducing the lifting factorization and the rounding operation into the CSMFB. Unlike the DTCWP, R-CSFMB can be designed directly as a M-channel system, and thus it can obtain better stopbband attenuation even if lifting coefficients are rounded. In the simulation, the directionality and the shift-invariance of the R-CSMFB are demonstrated. Moreover, the R-CSMFB is applied to image denoising as a practical application and shown its better performance than that of the R-DTCWT.

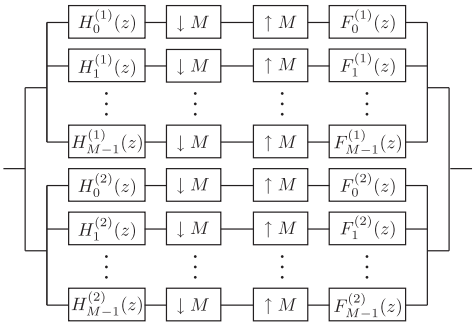


Fig. 1. M-channel DTCWTs and CSMFBs

Notations: **I** and **J** are the identity and the reversal identity matrices, repectively. $j := \sqrt{-1}$. $\overline{\alpha}$ is the conjugate of $\alpha \in \mathbb{C}$. $H(z)$, $H^*(z)$ and $\tilde{H}(z)$, $H(\omega)$ are defined as $H(z) := \sum_n h(n)z^{-n}$, $H^*(z) := \sum_n \overline{h(n)}z^{-n}$, and $\tilde{H}(z) := H^*(z^{-1})$. The $M \times L$ FB means the M channel filter bank with the filter length of L.

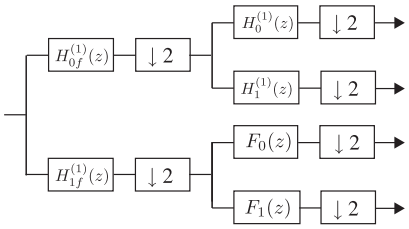
II. Review

A. Dual-Tree Complex Wavelet Transform and Dual-Tree Complex Wavelet Packet

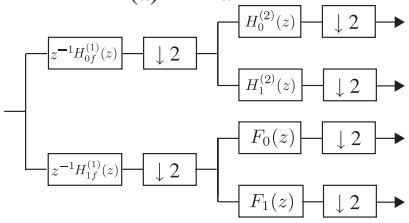
The DTCWT is constructed by two maximally decimated perfect reconstruction FBs, as illustrated in Fig. 1. In this case, the pair $\{H_k^{(1)}(z), H_k^{(2)}(z)\}$ $(0 \le k \le M-1)$ is designed to satisfy the following equations [1]:

$$H_k^{(2)}(e^{j\omega}) = e^{-j\theta_k(\omega)} H_k^{(1)}(e^{j\omega}), \tag{1}$$
$$\theta_0(\omega) = \left(d + \frac{1}{2}\right)(M-1)\omega - p\omega,$$

where $d \in \mathbb{Z}$ denotes the delay between the primal FB and the dual FB, $\forall p \in \left\{0, \ldots, \left\lceil \frac{M}{2} \right\rceil - 1\right\}, \forall \omega \in \left[\frac{2\pi}{M}p, \frac{2\pi}{M}(p+1)\right)$,



(a) Primal FB



(b) Dual FB

Fig. 2. Example of 2^r-band DTCWT ($r = 2$)

and

$$\theta_k(\omega) = \begin{cases} \frac{\pi}{2} - \left(d + \frac{1}{2}\right)\omega & \omega \in (0, 2\pi) \\ 0 & \omega = 0, \end{cases} \tag{2}$$

where $1 \le k \le M-1$.

Conventionally, the dual-tree complex wavelet packet (DTCWP) which is the design method of 2^r-channel DTCWT has been proposed [3]. Specifically, it is obtained by cascading a 2-channel DTCWT $\{H_k^{(1)}(z), H_k^{(2)}(z)\}_{k=0,1}$, and 2-channel perfect reconstruction FBs (PRFBs) $\{F_0(z), F_1(z)\}$ and $\{H_{0f}^{(1)}(z), H_{1f}^{(1)}(z)\}$, as shown in Fig. 2. The 2-channel DTCWT can be designed by the several methods, such as "q-shift solution" and "common factor solution" (spectral factorization) in [1]. The problem of the DTCWP is that good stopband attenuation cannot be obtained if the filter length of the whole system is restricted to be short. It is because the design degree of freedom for filter optimization remains only in the cascaded 2-channel FBs with the limited filter length.

B. Cosine-sine modulated filter banks

A CSMFB also consists of two maximally decimated filter banks as in Fig. 1. Its filter coefficients of $\{H_k^{(1)}(z), H_k^{(2)}(z)\}$ are expressed as follows [4]:

$$\begin{cases} h_k^{(1)}(n) := 2p(n)\cos\left((k+\frac{1}{2})\frac{\pi}{M}\left(n - \frac{N-1}{2}\right) + \theta_k\right) \\ h_k^{(2)}(n) := 2p(n)\sin\left((k+\frac{1}{2})\frac{\pi}{M}\left(n - \frac{N-1}{2}\right) + \theta_k\right), \end{cases} \tag{3}$$

where $k = 0, \ldots, M-1$, $\theta_k = (-1)^k\frac{\pi}{4}$ and $p(n)$ is the prototype filter. Let $\mathbf{E}^{(1)}(z)$ and $\mathbf{E}^{(2)}(z)$ be the polyphase matrices of the primal and the dual FBs, respectively, then, from (3), it follows that $\mathbf{E}^{(2)}(z) = \mathbf{\Gamma}\tilde{\mathbf{E}}^{(1)}(z)$, where $\mathbf{\Gamma}$ is the diagonal matrix $\mathbf{\Gamma} := diag((-1)^k)$ $(0 \le k \le M-1)$. Therefore, if the primal FB satisfies the perfect reconstruction (PR) property $\tilde{\mathbf{E}}^{(1)}(z)\mathbf{E}^{(1)}(z) = \mathbf{I}$, the dual FB $\mathbf{E}^{(2)}(z)$ system also satisfies the PR property. In our previous work, the directionality and the shift-invariance of CSMFBs are derived according to the relationship of cosine and sine modulation between two FBs [5].

III. Rational-valued Cosine Sine Modulated Filter Banks Based on the Lifting Factorization

In this section, the R-CSMFB is proposed. First, we consider a lattice strucutre corresponding to the CMFB. It is known as the extended lapped transform [6] expressed by:

$$\mathbf{E}^{(1)}(z) = \mathbf{C}_{\rm IV}\mathbf{W}\mathbf{\Lambda}(z)\mathbf{D}_0^{(1)}\prod_{k=1}^{K-1}\left(\mathbf{\Lambda}(z^2)\mathbf{D}_k^{(1)}\right)$$
$$\mathbf{D}_k^{(1)} = \begin{bmatrix} -\mathbf{C}_k & \mathbf{S}_k\mathbf{J} \\ \mathbf{J}\mathbf{S}_k\mathbf{J} & \mathbf{J}\mathbf{C}_k\mathbf{J} \end{bmatrix}, \mathbf{\Lambda}(z) = \begin{bmatrix} z^{-1}\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}, \tag{4}$$

where $\mathbf{C}_k = diag\left(\cos\theta_{k0}, \ldots, \cos\theta_{k\ M/2-1}\right)$, $\mathbf{S}_k = diag\left(\sin\theta_{k0}, \ldots, \sin\theta_{k\ M/2-1}\right)$ and $\mathbf{C}_{\rm IV}$ denotes the type-IV DCT. The structure is depicted in Fig. 3 (a). Then, we convert the lattice structure into lifting steps.

In general, the rotation matrix can be factorized into lifting steps as follows:

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} 1 & p \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ u & 1 \end{bmatrix}\begin{bmatrix} 1 & p \\ 0 & 1 \end{bmatrix}, \tag{5}$$

where $p = \frac{\cos\theta - 1}{\sin\theta}$, $u = \sin\theta$ are called lifting coefficients. The lifting factorization enables us to design rational-valued transformation simply by rounding lifting coefficients $round(p \times 2^N)/2^N$, $round(u \times 2^N)/2^N$, while preserving the perfect reconstruction condition. In fact, the 2-channel R-DTCWT is designed based on the lifting (or butterfly matrices) factorization [2].

As shown in Fig. 3 (b), each $\mathbf{D}_k^{(1)}$ can be reduced as

$$\mathbf{D}_k^{(1)} = \begin{bmatrix} -\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}\begin{bmatrix} \mathbf{I} & \mathbf{P}_k\mathbf{J} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{J}\mathbf{U}_k & \mathbf{I} \end{bmatrix}\begin{bmatrix} \mathbf{I} & \mathbf{P}_k\mathbf{J} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}, \tag{6}$$

where $\mathbf{P}_k = diag\left(p_{k0}, \ldots, p_{kM/2-1}\right)$ and $\mathbf{U}_k = diag\left(u_{k0}, \ldots, u_{kM/2-1}\right)$. Since $\mathbf{C}_{\rm IV}$ is constructed by some rotation and butterfly matrices [7] (described in Fig. 4), it can also be converted to the lifting-based factorization.

On the other hand, the lifting coefficients of $\mathrm{E}^{(2)}(z)$ can be determined immediately by using the relationship of $\mathbf{E}^{(2)}(z) = z^{-K}\mathbf{\Gamma}\tilde{\mathbf{E}}^{(1)}(z) = z^{-K}\mathbf{\Gamma}(\mathbf{E}^{(1)}(z))^{-1}$. Specifically, the lifting factorization of $\mathbf{E}^{(2)}(z)$ is expressed as follows:

$$\mathbf{E}^{(2)}(z) = \mathbf{\Gamma}\prod_{k=K-1}^{1}\left(\left(\mathbf{D}_k^{(1)}\right)^{-1} z^{-2}\mathbf{\Lambda}(z^{-2})\right) \times \left(\mathbf{D}_0^{(1)}\right)^{-1} z^{-1}\mathbf{\Lambda}(z^{-1})\mathbf{W}\mathbf{C}_{\rm IV} \tag{7}$$

According to the above expression, it can be verified that

$$\mathbf{D}_k^{(2)} = \left(\mathbf{D}_k^{(1)}\right)^{-1} = \begin{bmatrix} \mathbf{I} & -\mathbf{P}_k\mathbf{J} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{J}\mathbf{U}_k & \mathbf{I} \end{bmatrix}\begin{bmatrix} \mathbf{I} & -\mathbf{P}_k\mathbf{J} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}\begin{bmatrix} -\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}. \tag{8}$$

It indicates that the lifting coefficients of $\mathbf{E}^{(2)}(z)$ are just the sign-altered versions of those of $\mathbf{E}^{(1)}(z)$.

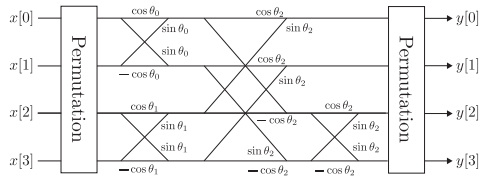


Fig. 4. The rotation-matrix-based structure of the type-IV DCT. (Example: 4×4, $\theta_0 = \pi/16$, $\theta_1 = 5\pi/16$, $\theta_2 = \pi/4$)

In [4], an another lattice structure of $\mathbf{E}^{(2)}(z)$ based on type IV DST has been presented. In that case, the lattice strucuture can be expressed by

$$\mathbf{E}^{(2)}(z) = \mathbf{S}_{\rm IV}\mathbf{W}\mathbf{\Lambda}(z)\mathbf{D}_0^{(2)}\prod_{k=1}^{K-1}\left(\mathbf{\Lambda}(z^2)\mathbf{D}_k^{(2)}\right)$$
$$\mathbf{D}_k^{(2)} = \begin{bmatrix}\mathbf{C}_k & \mathbf{S}_k\mathbf{J}\\ \mathbf{J}\mathbf{S}_k\mathbf{J} & -\mathbf{J}\mathbf{C}_k\mathbf{J}\end{bmatrix} = \begin{bmatrix}-\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}\mathbf{D}_k^{(1)}\begin{bmatrix}-\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}. \quad (9)$$

From the above equations and the relationship $\mathbf{S}_{\rm IV} = \mathbf{\Gamma}\mathbf{C}_{\rm IV}\mathbf{J}$, the lifting coefficients for $\mathbf{E}^{(1)}(z)$ can also be reused for $\mathbf{E}^{(2)}(z)$.

IV. Simulation

We design M-channel R-CSMFBs ($M > 2$) based on the lifting approach described in the previous section, and evaluate its directional selectivity andshift-invariance. For a comparison, M-channel R-DTCWT ($M > 2$) is designed by the DTCWP [3] cascading the 2-channel spectral-factorization-based R-DTCWT [2]. The cascaded 2-channel FBs $\{F_0(z), F_1(z)\}$ and $\{H_{0f}^{(1)}(z), H_{1f}^{(1)}(z)\}$ described in Sec. II-A are implemented by the lifting steps of the rotation matrices given in (5) and the delay matrices. Both M-channel transforms are optimized to minimize the stopband attenuation Δ specified by

$$\Delta = \sum_{k=0}^{M-1}\int_{\Omega_s(k)} |H_k^{(1)}(e^{j\omega})|^2 d\omega, \quad (10)$$

where $\Omega_s(k)$ is stoppband of the $H_k^{(1)}(e^{j\omega})$, $(0 \le k \le M-1)$. In this paper, we set as $\Omega_s(k) = [0, \frac{\pi}{M}(k-2)] \cup [\frac{\pi}{M}(k+2), \pi]$. If $\frac{\pi}{M}(k-2) \le 0$ or $\frac{\pi}{M}(k+2) \ge \pi$, the corresponding interval is ignored.

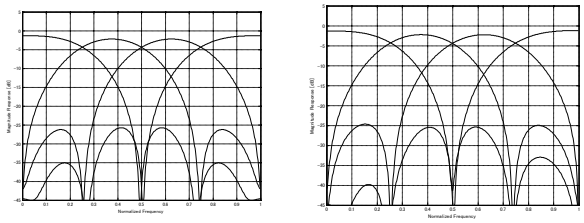
A. Design Example

The design examples of the R-CSMFB and the R-DTCWT are given in this section. First, the R-CSMFB and the R-DTCWT are designed as a real-valued FB minimizing the stopband attenuation. Then, the real-valued lifting coefficients are rounded to the rational coefficients from 8 to 4 bit word length. The results of the stopband attenuation Δ are listed in Table I. The numbers of the filter length of the cascaded 2-channel PUFBs for the 4×16 (# of channel $\times$ # of filter length) and 8×22 R-DTCWTs and the 4×22 and 8×36 R-DTCWTs are 4 and 6, respectively. Thus, the design degree of freedom for optimization is quite limited. On the other hand, since the CSMFBs with multiple channels can be designed directly, more design degree of freedom can be obtained, and thus it can achieve better stopband attenuation.

TABLE I
Stopband attenuation

4-channel R-DTCWT					
Word length	8 bit	7 bit	6 bit	5 bit	4 bit
Δ (L=16)	2.2315	2.2857	2.5866	3.4617	7.8892
Δ (L=22)	0.6356	0.5374	1.6500	1.3035	3.4647
4-channel R-CSMFB					
Word length	8 bit	7 bit	6 bit	5 bit	4 bit
Δ (L=8)	0.2477	0.2278	0.3664	0.3069	1.5346
Δ (L=16)	0.3863	0.8409	0.9212	1.4547	1.7782

8-channel R-DTCWT					
Word length	8 bit	7 bit	6 bit	5 bit	4 bit
Δ (L=22)	135.0658	135.0658	134.1306	143.3687	134.6823
Δ (L=36)	72.5471	71.3914	76.2479	77.1526	75.1663
8-channel R-CSMFB					
Word length	8 bit	7 bit	6 bit	5 bit	4 bit
Δ (L=16)	2.2984	3.0656	36.1532	31.5684	42.0679
Δ (L=32)	1.3973	1.5174	1.5949	3.2597	11.3918



(a) Real (b) Rational

Fig. 5. Frequency responses of the designed 4×8 CSMFB. (a): Real-valued CSMFB, (b) Rational-valued CSMFB with 5 bit lifting coefficients.

Fig. 5 depicts the frequency responses of the designed 4×8 CSMFB and R-CSMFB. It can be observed that the frequency responses of R-CSMFB in Fig. 5(b) are robustly approximating those of the original CSMFB given in Fig. 5(a) under rounding operation.

B. Directional Selectivity

The 2-D wavelet basis functions obtained by the 4-channel R-CSMFB are shown in Fig. 6. Clearly, the basis functions are oriented, which indicates the R-CSMFB achieves rich directional selectivity under rounding operation.

C. Evaluation of Shift Invariance

In this section, shift invariance of the R-CSMFB and the R-DTCWT (4×16, 4 bit word length) are evaluated. First, we apply the R-CSMFB and the R-DTCWT to the input impulse and its shifted impulse independently. Then, compute the correlation of the output subband signals. Specifically, let $s_{k,r}^{(1)}$ and $s_{k,r}^{(2)}$ $(0 \le k \le 3)$ be the k-th subband signals of the first and the second system which are obtained from the r-sample $(0 \le r \le 3)$ shifted original signal (the original one is denoted by $r = 0$). Then, we evaluate the shift invariant property by the normalized correlation of $s_{k,r}^{(1)} + s_{k,r}^{(2)}$ and $s_{k,0}^{(1)} + s_{k,0}^{(2)}$ denoted by $\Phi(k,r)$ $(0 \le \Phi(k,r) \le 1)$. Table II shows the averaged correlation values $\Phi_A(k) = \frac{1}{4}\sum_{r=0}^{3}\Phi(k,r)$. As

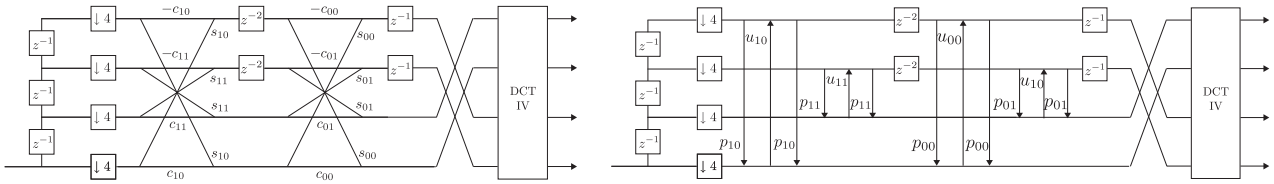


Fig. 3. The structure of the 4 channel CMFBs. (Left: rotation-matrix-based structure, Right: lifting-based structure)

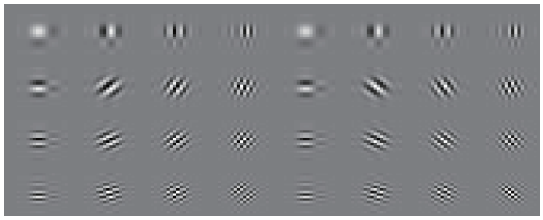


Fig. 6. 32 directional basis functions with size of 16×16 obtained by the 4×16 CSMFB whose lifting coefficents are rounded to 4 bit word length.

TABLE II
CORRELATION RESULTS

4-Channel R-DTCWT (filter length: 16)				
Channel	$k=0$	$k=1$	$k=2$	$k=3$
$\Phi_A(k)$	0.9505	0.9252	0.9477	0.9124
4-Channel R-CSMFB (filter length: 16)				
Channel	$k=0$	$k=1$	$k=2$	$k=3$
$\Phi_A(k)$	0.9962	0.9940	0.9948	0.9935

shown in Table II, the R-CSMFB satisfies the better shift invariant property than the R-DTCWT.

D. Image Denoising

In this simulation, image denoising is demonstrated to verify the performance of R-CSMFB. The gaussian random noise with variance σ^2 $(\sigma = 15, 20, 25, 30)$ is added to the test images *Lena* and *Barbara*. The noisy images are first decomposed by 2-level 8-channel of the R-CSMFB and the R-DTCWT, then performed by hard-thresholding with 3σ. The lifting coefficients of both transfomations are allocated 8 bit word length. For the numerical metric of the denoising performance, we use PSNR. As shown in Table III and Fig. 7, R-CSMFB shows better visual quality than those of the R-DTCWT, thanks to the better stopband attenuation of the R-CSMFB.

V. CONCLUSION

In this paper we proposed the directional and shift-invariant transform based on M-channel R-CSMFB by the lifting factorization and the rounding operation. It is clarified that the lifting coefficients for the primal FB can be reused for the dual FB. The R-CSMFB provides not only better stopband attenuation than the R-DTCWT, but also the properties of shift invariance and rich directional selectivity under the rational value constraint. Finally, the R-CSMFB is applied to image denoising as one of the application, and shown its better performance than that of the R-DTCWT.

TABLE III
DENOISING RESULTS

Lena R-DTCWT (2 level, filter length: 22)				
σ	$\sigma=15$	$\sigma=20$	$\sigma=25$	$\sigma=30$
PSNR [dB]	30.4668	29.0718	28.0236	27.1358
Lena R-CSMFB (2 level, filter length: 16)				
σ	$\sigma=15$	$\sigma=20$	$\sigma=25$	$\sigma=30$
PSNR [dB]	31.0437	29.6159	28.4847	27.5970
Barbara R-DTCWT (2 level, filter length: 22)				
σ	$\sigma=15$	$\sigma=20$	$\sigma=25$	$\sigma=30$
PSNR [dB]	28.8703	27.3681	26.1581	25.2805
Barbara R-CSMFB (2 level, filter length: 16)				
σ	$\sigma=15$	$\sigma=20$	$\sigma=25$	$\sigma=30$
PSNR [dB]	29.8567	28.2664	26.9738	26.0463

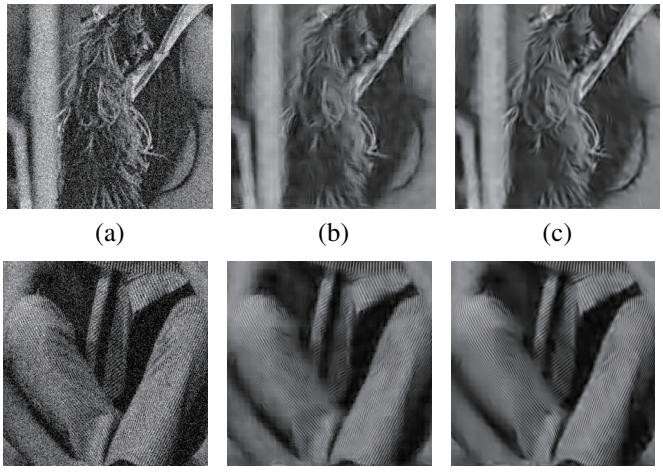


Fig. 7. Denoising results. (a) and (d): Noisy images of *Lena* and *Barbara* ($sigma = 30$) (b) and (e): R-DTCWT (8×22) (c) and (f): R-CSMFB (8×16)

REFERENCES

[1] I. W. Selesnick, R. G. Baraniuk and N. C. Kingsbury, "The dual-tree complex wavelet transform," *IEEE Signal Process. Magazine*, vol. 22 pp. 123–151, Nov. 2005.

[2] A. Abbas and Trac D. Tran, "Rational coefficient dual-tree complex wavelet transform: design and implementation," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3523–3534, Aug. 2008.

[3] I. Bayram and I. W. Selesnick, "On the dual-tree complex wavelet packet and M-band transform," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2298–2310, June 2008.

[4] A. Viholainen, J. Alhava and M. Renfors, "Implementation of parallel cosine and sine modulated filter banks for equalized transmultiplexer systems," in *Proc. ICASSP*, vol. 6, pp. 3625–3628, June 1984.

[5] S. Kyochi, T. Uto, M. Ikehara, "Dual-tree complex wavelet transform arising from cosine-sine modulated filter banks," in *Proc. ISCAS*, pp.2189-2192, May 2009.

[6] S. Malvar, "Extended lapped transforms: properties, applications and fast algorithms," *IEEE Trans. Signal Process.*, vol. 40, no. 11, pp. 2703–2714, June 1992.

[7] Z. Wang, "Fast algorithms for the discrete W transform and for the discrete fourier transform," *IEEE Trans. Signal Process.*, vol. 40, no. 11, pp. 803–816, June 1984.