Multiscale Image Fusion Using the Undecimated Wavelet Transform With Non-Orthogonal Filter Banks

Andreas Ellmauthaler, Eduardo A. B. da Silva, Carla L. Pagliari and Marcelo M. Perez

***Abstract*— In this work a pixel-level multiscale image fusion framework is introduced which utilizes a new class of non-orthogonal filter banks. These filter banks were especially designed for the purpose of image fusion and offer some useful properties such as a short support size of the analysis filter pair whilst preserving the regularity of the synthesis filters. As a result, the unwanted spreading of coefficient values around overlapping image singularities, usually complicating the feature selection process, is minimized. Furthermore, the introduction of blocking artifacts in the fused reconstruction is avoided. In the course of this work we will show that the combination of the Undecimated Wavelet Transform with such filter banks leads to a fusion framework which is able to significantly outperform traditional multiscale fusion approaches for a large group of images, derived from different sensor modalities.**

***Keywords*— Image Fusion, Undecimated Wavelet Transform, Non-Orthogonal Filter Banks**

***Resumo*— Neste trabalho, é introduzido um método de fusão de imagens usando técnicas multiescalas ao nível do pixel, utilizando uma nova classe de bancos de filtros não-ortogonais. Estes bancos de filtros foram especialmente desenvolvidos para o propósito de fusão de imagens e oferecem propriedades úteis, tais como suporte curto dos filtros de análise e preservação da regularidade dos filtros de síntese. Assim, o espalhamento indesejável dos coeficientes na vizinhança de singularidades sobrepostas, que geralmente dificultam o processo de seleção de *features*, é minimizado. Além disso, é evitada a introdução de artefatos de blocagem na reconstrução fundida. Ao longo deste trabalho, mostramos que a combinação da Transformada Wavelet Não-Decimada com tais bancos de filtros é capaz de oferecer desempenho superior às abordagens tradicionais de fusão de imagens usando técnicas multiescalas para um grande grupo de imagens, derivadas de diferentes modelos de sensores.**

***Palavras-Chave*— Fusão de Imagens, Transformada Wavelets Não-Decimada, Bancos de Filtros Não-Ortogonais**

I. INTRODUCTION

Within the last decades substantial progress was achieved in the imagery sensor field. Improved robustness and increased resolution of modern imaging sensors together with cheap fabrication costs have made the use of multiple sensors common in a wide range of imaging applications. Often it is convenient to merge such multisensor data into one composite representation for interpretation purposes. In image-based applications this plethora of combination techniques became generally known as image fusion and is nowadays a promising research area.

Image fusion can be summarized as the process of integrating complementary and redundant information from multiple images into one composite image that contains a 'better' description of the underlying scene than any of the individual source images. Hence, the fused image should be more useful for visual inspection or further machine processing [1]. However, fusing images is often not a trivial process, since: a) the source images may come from different types of sensors (e.g. with different dynamic ranges and resolutions); b) they tend to exhibit complementary information (e.g. features which appear in some source images but not in all) or c) they may show common information but with reversed contrast, which complicates the fusion process significantly. Furthermore, a fusion approach which is independent of a-priori information about the inputs and produces a composite image which appears 'natural' to a human interpreter is highly desirable [2].

The majority of image fusion strategies can be classified into pixel- and region-level approaches. As for pixel-level image fusion techniques, each pixel in the fused image is determined by a set of pixels in the source images. Region-level fusion approaches typically segment the images into regions and perform the fusion based on the extracted regions. Usually, the latter scheme has advantages over pixel-based techniques since some drawbacks, such as blurring effects, high sensitivity to noise and misregistration, can be avoided [2]. However, the final fusion performance of region-based image fusion methods highly depends on the quality of the segmentation process. In other words, segmentation errors such as under- or over-segmentation may lead to the absence or degradation of certain features in the fused image [3]. Please note that only pixel-level image fusion will be addressed in the course of this work.

In the literature plenty of pixel-level fusion methods can be found, reaching from simple averaging techniques to more complex methods employing multiscale (MS) transforms. As for the latter group, the actual fusion takes place after transforming each source image into the transform domain (e.g. wavelet domain). The final fused image is obtained by taking the inverse transform of the composite representation. The set of used MS transforms includes Pyramid Transforms, the Discrete Wavelet Transform (DWT), the Undecimated Wavelet Transform (UWT), the Dual-Tree Complex Wavelet Transform (DTCWT), the Curvelet Transform, the Contourlet Transform and the Nonsubsampled Contourlet Transform (NSCT).

Recent research efforts showed that the best MS pixel-level fusion results are achieved using the NSCT, followed by the DTCWT and the SWT [4]. In this work, we demonstrate that for images derived from different sensor modalities the NSCT and the DTCWT can be significantly outperformed by an efficient and easy-to-implement method consisting of the UWT in combination with a new class of non-orthogonal filter banks. The used filters were especially designed for the purpose of image fusion and exhibit some useful properties such as a very short support size of the analysis filter pair whilst preserving the regularity of the synthesis filters. Furthermore, we will

Andreas Ellmauthaler and Eduardo A. B. da Silva are with the Universidade Federal do Rio de Janeiro, PEE/COPPE/DEL, Rio de Janeiro, Brazil, E-mail: {andreas.ellmauthaler, eduardo}@lps.ufrj.br; Carla L. Pagliari is with the Instituto Militar de Engenharia, Department of Electrical Engineering, Rio de Janeiro, Brazil, E-mail: carla@ime.eb.br; Marcelo M. Perez is with the Exército Brasileiro, Departamento de Ciência e Tecnologia, E-mail: perez@ime.eb.br The authors would like to thank the Brazilian Funding Agencies CAPES (Pro-Defesa) and FINEP, and the Army Technological Center - CTEx for their financial support.

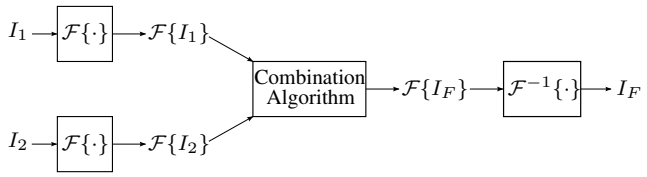

Fig. 1. Generic fusion scheme in the transform domain.

see that the removal of the orthogonality restriction allows for the implementation of filter banks such that both synthesis filters are positive. As a result, we successfully reduce the coefficient spreading problem that usually complicates the feature selection process, thus avoiding the introduction of artifacts during reconstruction.

The structure of this paper is as follows: Section II gives an overview on MS pixel-level image fusion, whereas Section III is dedicated to the introduction of the proposed image fusion framework. The obtained results are presented and compared with other state-of-the-art fusion frameworks in Section IV, before we state our main conclusions in Section V.

II. Multiscale Image Fusion

In the last two decades pixel-level image fusion gained considerable attention from the image processing community. A straightforward approach for pixel-level fusion is to take the pixel-by-pixel average of the source images. Although this method is very simple to implement, it presents several drawbacks including reduced contrast. This can lead to a severe loss of information, which, in general, renders the averaging method unsuitable for most fusion scenarios.

In general, pixel-level techniques can be divided into spatial and transform domain techniques. In the spatial domain, the fusion is performed by combining all input images in a linear or non-linear fashion. Assuming that $g(\cdot)$ represents a function which governs the combination of the input images $I_k, k = 1, \ldots, K$, commonly known as the "fusion rule", spatial domain techniques can be defined as [5]

$$I_F[m,n] = g(I_1[m,n], \ldots, I_K[m,n]) \ , \qquad (1)$$

where m, n represents the spatial location in the input images and the fused image I_F.

Transform domain techniques map (transform) each source image into the transform domain, where the actual fusion process takes place. The final fused image is obtained by taking the inverse transform of the composite representation. The main motivation behind moving to the transfer domain is to work within a framework where the image's salient features are more clearly depicted than in the spatial domain. If we let $\mathcal{F}\{\cdot\}$ and $\mathcal{F}^{-1}\{\cdot\}$ represent the forward and inverse transforms, respectively, transform domain techniques can be defined as [5]

$$I_F[m,n] = \mathcal{F}^{-1}\{g(\mathcal{F}\{I_1[m,n]\}, \ldots, \mathcal{F}\{I_K[m,n]\})\} \ . \quad (2)$$

Fig. 1 illustrates this process for two input images. While many different techniques have been proposed, most of the more successful transform domain approaches use MS transforms [5]. This is motivated by the fact that images tend to present features in many different scales. In addition, the human visual system seems to exhibit high similarities with the properties of MS transforms. More precisely, strong evidence exists that the entire human visual field is covered by neurons that are selective to a limited range of orientations and spatial frequencies, and can detect local features like edges and lines. This makes them very similar to the basis functions of MS transforms [6]. Thus, in this work we opted to address only MS pixel-level image fusion.

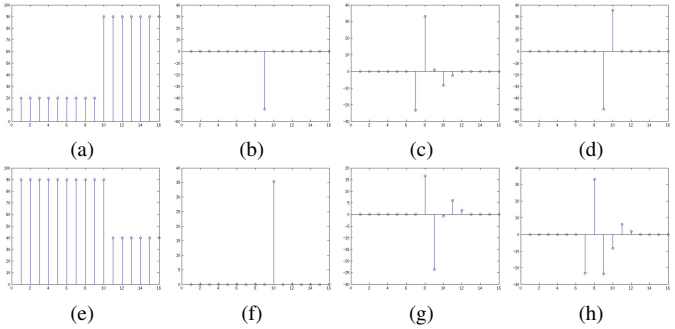

Fig. 2. Coefficient spreading effect. (a) and (e) Input signals. (b) and (f) Haar filtered input signals. (c) and (g) 'db3' filtered input signals. (d) Fusion of the Haar filtered signals. (h) Fusion of the 'db3' filtered signals.

So far in this section we have made no restrictions on the MS transform in use. As stated in various studies (e.g. [1], [4]), the shift-invariance property, provided by the Undecimated Wavelet Transform (UWT), by the Nonsubsampled Contourlet Transform (NSCT) and approximately by the Dual-Tree Complex Wavelet Transform (DTCWT) is a highly desirable property in image fusion applications. Furthermore, due to the redundancy involved, these transforms are more robust to discontinuities introduced in the sub-band images, leading to fused images which are less prone to reconstruction errors. Motivated by these observations, we will therefore discard critically-sampled transforms such as the DWT, the Curvelet Transform and the Contourlet Transform and solely focus on redundant transforms in our ongoing discussion.

Often the fusion performance of traditional MS pixel-level approaches deteriorates when dealing with input images exhibiting adjacent or locally overlapping features. This is mainly due to the support of the filters used during the decomposition process which results in an undesirable spreading of coefficient values over the neighborhood of salient features, possibly leading to conflicting information being associated with the same corresponding coefficients in the transform domain. This considerably complicates the feature selection process and may lead to the introduction of distortions which can be perceived as ringing artifacts or a substantial loss of information in the fused image.

Fig. 2 attempts to illustrate the impact of the length of the chosen filter bank on the fusion performance. In this example the high-pass portions of two 1-D step functions are fused using one stage of the 2-tap Haar and 6-tap 'db3' filters. The applied fusion rule is the so-called 'choose max' fusion rule as given in eq. (4), where the coefficient yielding the highest energy is directly transferred to the fused signal. The high-pass sub-bands, obtained by applying the Haar filter, can be seen in Figs. 2(b) and (f), whereas the result using the 'db3' filter is illustrated in Figs. 2(c) and (g). It can be observed that the 6-tap 'db3' filter needs five coefficients to represent the step change. Thus, although most energy is concentrated in the central coefficient, the remaining four coefficients correspond to regions where no change in the signal value occurred. When attempting to fuse the two 'db3' filtered high-pass sub-bands we are confronted with a problem which can not be resolved, namely, to combine the two signals without losing information. This can be observed in Fig. 2(h), where not all non-zero coefficients from Figs. 2(c) and (g) could be incorporated. On the other hand, the Haar filtered signal contains only one non-zero coefficient corresponding exactly to the position of the signal transition. Thus, as illustrated in Fig. 2(d), both non-

zero coefficients are transferred to the fused image without any loss of information. Even though the situation depicted in Fig. 2 may seem at first somewhat artificial, we will see that multisensor imagery often exhibits similar properties. Hence, for these images the fusion performance considerably degrades with an increase of the filter size.

We can therefore address the problem of choosing a proper redundant MS transform by looking at its ability to incorporate a filter bank with a sufficiently small support size, thus minimizing the coefficient spreading problem. From this point of view, the UWT appears to be an attractive choice. This is so because, due to the standard tensor product construction in two dimensions, the UWT offers directionality without increasing the overall length of the implemented filter bank - a property not shared by the NSCT and the DTCWT. As for the NSCT, the increased filter lengths are mainly due to the iterated nature of the nonsubsampled directional filter bank involved. In particular, for every increase in number of directions by a power of two, another filter bank level needs to be added (see [7] for a thorough discussion of the NSCT). Thus, the combined support of the filters within one particular filter bank branch is equivalent to the one of the convolution of all individual filters within the respective branch. In the case of the DTCWT, as reported in [8], the increased filter length is due to the half-sample delay condition imposed on the filter banks involved, which results in longer filters than in the real wavelet transform case. From the above, we can conclude that, even though the NSCT and the DTCWT posses some very useful properties, such as their ability to incorporate a higher number of orientations, they are, in general, less suited to implement filter banks with a small support size.

In this paper we attempt to show that very good fusion results for infrared-visible and medical image pairs can be obtained by simply applying the UWT in combination with a new class of non-orthogonal filter banks. More specifically, we will demonstrate that the nonsubsampled nature of the UWT allows for the implementation of filter banks which are able to minimize both the coefficient spreading problem and the introduction of artifacts in the final fused image. This leads to a fusion framework which provides clear advantages over traditional MS approaches.

III. The UWT-based fusion scheme with non-orthogonal filter banks

An input image can be represented in the UWT domain by a sequence of detail images at different scales and orientations along with an approximation image at the coarsest scale. Hence, the UWT decomposition of an input image I_k can be represented as

$$y_k = \{y_k^1, y_k^2, \ldots, y_k^J, x_k^J\} \ , \tag{3}$$

where x_k^J stands for the approximation image at the lowest scale J and $y_k^j, j = 1, \ldots, J$ represent the detail images at level j. Each of these is comprised of three orientation bands $y_k^j = \{y_k^j[\cdot, 1], y_k^j[\cdot, 2], y_k^j[\cdot, 3]\}$. In order to simplify the discussion, we assume, without loss of generality, that the fused image will be generated from two source images I_A and I_B which are assumed to be registered prior to the fusion process.

As for the combination of the input image pair, a simple 'choose max' fusion rule will be used for all decomposed detail images. By this rule the coefficient yielding the highest energy is directly transferred to the fused decomposed representation. Hence, the fused detail images y_F^j are defined as

$$y_F^j[\mathbf{n}, p] = \begin{cases} y_A^j[\mathbf{n}, p] & \text{if } \left|y_A^j[\mathbf{n}, p]\right| > \left|y_B^j[\mathbf{n}, p]\right| \\ y_B^j[\mathbf{n}, p] & \text{otherwise} \end{cases} \ , \tag{4}$$

where the vector $\mathbf{n} = [m, n]$ represents the spatial location in a given orientation band p at decomposition level j. This choice is motivated by the fact that salient features, such as edges, lines or other discontinuities, result in large magnitude coefficients, and can be efficiently captured using the above mentioned fusion scheme.

Since high magnitudes in the low-pass approximation images do not necessarily correspond to important features within the source images, the approximation images will be treated differently. In our approach, the composite approximation coefficients are obtained by a simple averaging operation

$$x_F^J[\mathbf{n}] = \frac{x_A^J[\mathbf{n}] + x_B^J[\mathbf{n}]}{2} \ . \tag{5}$$

In the literature more sophisticated fusion rules can be found. However, since the main contribution of the proposed fusion framework is focused on the employed non-orthogonal filter banks rather than on the used fusion rule, the 'choose max' rule in combination with an averaging of the approximation images will suffice for our purposes. Additionally, we believe that the conclusions drawn in this paper can also be applied to other, more powerful fusion rules. An exhaustive study on the influence of different coefficient merging techniques can be found in [9].

A. Filter Bank Design

Due to the nonsubsampled nature of the UWT, many ways exist to construct the fused image from its wavelet coefficients. For a given analysis filter bank (h, g), any synthesis filter bank $(\tilde{h}, \tilde{g})$ satisfying the perfect reconstruction condition

$$h[n] * \tilde{h}[n] + g[n] * \tilde{g}[n] = \delta[n] \tag{6}$$

can be used for reconstruction, where $\delta[n]$ represents an impulse at $n = 0$ [10]. This is considerably simpler and offers more design freedom than in the decimated case, where an additional anti-aliasing condition has to be obeyed, imposing substantial constraints on the filter bank design. In the remainder of this section we show how such filters can be designed for the purpose of image fusion. We would like to point out that none of these filters obey the anti-aliasing condition and can therefore only be used in the undecimated case.

We start our discussion by noting that for any pair of even symmetric analysis filters (h, g), the high-pass analysis filter g can be defined such that $g = \delta - h$. In this case, a simple choice for the synthesis filters is to define them as all-pass filters with $\tilde{h} = \tilde{g} = \delta$ which implies that the reconstruction of the original image I is obtained by co-addition of all detail images to the approximation image, that is

$$I[\mathbf{n}] = x^J[\mathbf{n}] + \sum_{j=1}^{J} \sum_{p=1}^{3} y^j[\mathbf{n}, p] \ . \tag{7}$$

This approach is frequently used in multispectral image fusion (e.g. fusion of high-resolution panchromatic images with low-resolution multispectral images) and has some interesting characteristics. For example, due to the lack of convolutions during reconstruction, no additional distortions are introduced when constructing the fused image. Furthermore, since the

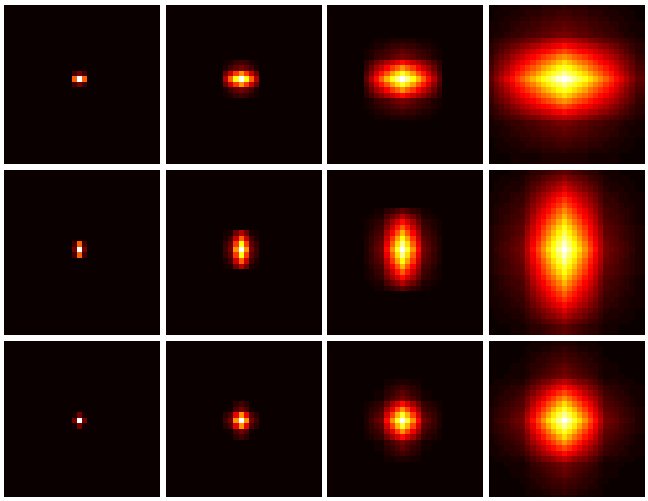

Fig. 3. Backprojection of a single wavelet coefficient at different scales and directions for the filter bank given in eq. (9). From top to bottom, the coefficient belongs to the horizontal, vertical and diagonal bands. From left to right, the scale increases from one to four. Each scale and direction has been normalized such that it occupies the full dynamic range.

fused image is obtained by a simple co-addition of all detail images and the approximation image, a very fast reconstruction is possible. As for the choice of h we use a B-spline filter. Consequently, the overall filter bank may be defined such that

$$\begin{array}{lcll} h[n] & & = & \dfrac{[1,2,1]}{4} \\ g[n] & = & \delta[n]-h[n] & = & \dfrac{[-1,2,-1]}{4} \\ \tilde{h}[n] & = & \tilde{g}[n] & = & [0,1,0] \end{array} , \qquad (8)$$

where h is derived from the B_1-spline function. In general B-spline functions have some remarkable properties which make them very good choices for wavelet analysis. For example, if we recall that the low-pass filter h in the z-transform domain can be factorized as $H(z) = (1+z^{-1})^L P(z)$, it can be shown that the B-spline function of degree $n = L-1$ is the shortest and most regular scaling function of order L, with $P(z) = 1$ [11].

However, the approach presented above comes with one disadvantage. More specifically, due to the employed all-pass filters, distortions introduced during the fusion process remain unfiltered in the reconstructed image. Hence, in order to improve the overall performance, we choose h and g as in eq. (8) but define the synthesis low-pass filter $\tilde{h}$ as h. In compliance with the perfect reconstruction condition of eq. (6) this yields that $\tilde{g} = \delta + h$, resulting in the filter bank

$$\begin{array}{lcll} h[n] & = & \tilde{h}[n] & = & \dfrac{[1,2,1]}{4} \\ g[n] & = & \delta[n]-h[n] & = & \dfrac{[-1,2,-1]}{4} \\ \tilde{g}[n] & = & \delta[n]+h[n] & = & \dfrac{[1,6,1]}{4} \end{array} . \qquad (9)$$

In this scenario $\tilde{g}$ consists entirely of positive coefficients, being thus no longer related to a wavelet function. On the other hand, such a lack of oscillations provides a reconstruction less prone to ringing artifacts. Additionally, unlike eq. (8), distortions introduced during the fusion stage are not transferred unprocessed to the reconstructed image. Fig. 3 shows the backprojection of a wavelet coefficient at different scales and directions for the filter bank given in eq. (9). Note that all images only exhibit positive values.

In Section II we discussed how different filter lengths can influence the fusion result. More specifically, we argued that filters with long support sizes lead to an unwanted spreading of coefficient values around image singularities, complicating the feature selection process. Following this reasoning, we propose a final modification of the filter bank given in eq. (9). Note that the analysis filters h and g can be factorized such that

$$\begin{array}{lcll} h[n] & = & \dfrac{[1,2,1]}{4} & = & \dfrac{[1,1]}{2} * \dfrac{[1,1]}{2} \\ g[n] & = & \dfrac{[-1,2,-1]}{4} & = & \dfrac{[1,-1]}{2} * \dfrac{[-1,1]}{2} \end{array} . \qquad (10)$$

Hence, we can move the 2^{nd} factor of h and g, respectively, to the synthesis filter pair without corrupting the perfect reconstruction condition. This yields the following filter bank:

$$\begin{array}{lcllcl} h[n] & = & \dfrac{[1,1]}{2} & \tilde{h}[n] & = & \dfrac{[1,3,3,1]}{8} \\ g[n] & = & \dfrac{[1,-1]}{2} & \tilde{g}[n] & = & \dfrac{[-1,-5,5,1]}{8} \end{array} . \qquad (11)$$

In this setup the analysis filter bank (h, g) consists of the very short 2-tap Haar filter pair whereas the synthesis filters correspond to a filter pair with larger support size and higher regularity. In other words, by employing this filter bank, the coefficient spreading problem (related to the support of the analysis filters) as well as the introduction of blocking artifacts during reconstruction (related to the regularity of the synthesis filters) is successfully reduced. We will see in the next section that this filter bank is particularly well suited for the fusion of infrared-visible and medical images, which tend to exhibit a high degree of overlapping salient information.

IV. Results

The performance of the proposed image fusion scheme was compared to the fusion results obtained by applying the Nonsubsampled Contourlet Transform (NSCT) and the Dual-Tree Complex Wavelet Transform (DTCWT). As for the NSCT and the DTCWT, we followed the recommendations published in [4] regarding the filter choices and (in case of the NSCT) number of directions. In the case of the UWT-based image fusion scheme, we utilized the filter banks of Section III-A. Hence, in our experiments we used the non-orthogonal filter banks of eqs. (8), (9) and (11). In order to avoid referring to filter banks by their respective equation numbers, we will associate names to them. Henceforth, the filter banks presented in eqs. (8), (9) and (11) will be referred to as 'Spline_1', 'Spline_2' and 'Spline_3' filter banks, respectively. Four decomposition levels were chosen for all transforms.

We performed simulations for 10 infrared-visible image pairs using the fusion rules given in eqs. (4) and (5). As for the objective evaluation of the achieved results, we use in this work three of the most widely used fusion metrics, namely, the performance measures proposed by Xydeas and Petrović $Q_{AB/F}$ [12], Piella Q_P [2] as well as the Mutual Information (MI), first introduced by Qu et al. [13] in the context of image fusion.

Tables I and II list the average results for all infrared-visible and medical image pairs, respectively. It can be noticed that the UWT in combination with the proposed new class of non-orthogonal filter banks significantly outperforms traditional fusion methods based on the NSCT and the DTCWT for all tested fusion metrics. More specifically, by looking at Tables I and II it can be noted that the best performance is achieved for the 'Spline_3' filter bank, followed by the 'Spline_2' filter bank (note that, for example, the results for the 'Spline_3'

TABLE I
FUSION RESULTS FOR INFRARED-VISIBLE IMAGE PAIRS.

Transform/Filter Bank	$Q_{AB/F}$	MI	Q_P
DTCWT	0.5664	0.1538	0.7707
NSCT	0.5786	0.1563	0.7719
Spline_1	0.5622	0.1570	0.7517
Spline_2	0.5791	0.1557	**0.7758**
Spline_3	**0.5946**	**0.1574**	0.7750

TABLE II
FUSION RESULTS FOR MEDICAL IMAGE PAIRS.

Transform/Filter Bank	$Q_{AB/F}$	MI	Q_P
DTCWT	0.6314	0.3853	0.6618
NSCT	0.6624	0.4035	0.6667
Spline_1	0.6519	0.4189	0.6552
Spline_2	0.6802	0.4229	**0.6845**
Spline_3	**0.7104**	**0.4288**	0.6687

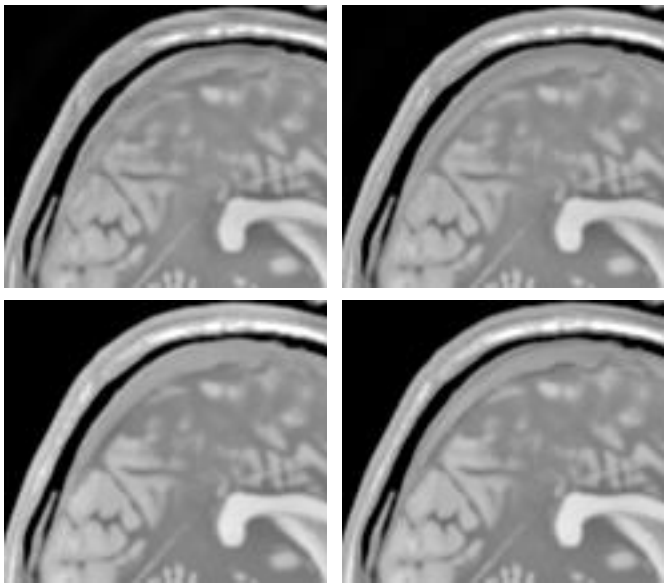

Fig. 4. Magnified fusion results of a medical image pair. (Top-left) DTCWT fused. (Top-right) NSCT fused. (Bottom-left) UWT fused with 'Spline_2' filter bank. (Bottom-right) UWT fused with 'Spline_3' filter bank.

filter bank are in all cases superior to the ones of DTCWT and NSCT). This is most evident when looking at the magnified fusion results of a medical image pair, depicted in Fig. 4. It can be seen that the DTCWT- and NSCT-based fusion schemes suffer from a significant loss of edge information, particularly noticeable at the outermost borders of the fused images (top-left and top-right corner of Fig. 4, respectively). There, information belonging to the skull bone (white stripe enclosed within the gray, tube-like structure) partially disappeared. This is due to the superposition of the skull bones, originating from the pair of source images, resulting in coefficient overlaps in the DTCWT and NSCT transform domain, which cannot be resolved by the fusion algorithm. As for the fusion results obtained by the UWT, this effect is reduced and the edge information is preserved to a much higher degree. Moreover, in case of the 'Spline_3' filter bank, the edges appear to be more accentuated than in the fusion scenario using the 'Spline_2' filter bank, further indicating the perceptual superiority of the proposed fusion approach. Please note that no performance gains could be achieved for the 'Spline_1' filter bank. As mentioned in Section III-A, this is mainly due to the absence of filtering operations during reconstruction, since in this case distortions introduced during the fusion process remain unaltered in the fused image.

The MATLAB implementation of the proposed UWT-based image fusion framework with non-orthogonal filter banks, together with the fusion results for all tested images, is available for download at http://www.lps.ufrj.br/profs/eduardo/fusion/SBrT_2012.

V. CONCLUSIONS

In this paper a novel UWT-based image fusion framework is introduced. The proposed approach takes advantage of the increased filter design freedom of the UWT which allows for the development of a new class of non-orthogonal filter banks. The filters were especially designed for the purpose of image fusion and exhibit useful properties such as a short support size of the analysis filter pair whilst preserving the regularity of the synthesis filters. Thereby, the unwanted spreading of coefficient values around overlapping image singularities is successfully reduced and the introduction of artifacts in the fused reconstruction is avoided. Moreover, the nonsubsampled nature of the UWT permits the design of filter banks where both synthesis filters exhibit only positive coefficients. Such filters provide a reconstructed, fused image less vulnerable to ringing artifacts. The results obtained using objective metrics indicate that our solution leads to a fusion framework which provides clear advantages over traditional multiscale fusion approaches based on state-of-the-art transforms such as the DTCWT and the NSCT. Additionally, the perceptual superiority of the proposed framework was suggested by informal visual inspection of a fused medical image pair.

REFERENCES

[1] Z. Zhang and R. S. Blum, "A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application," *Proceedings of the IEEE*, vol. 87, no. 8, pp. 1315–1326, August 1999.

[2] G. Piella, *Adaptive Wavelets and their Applications to Image Fusion and Compression*, Ph.D Thesis, University of Amsterdam, Amsterdam, Netherlands, 2003.

[3] A. Ellmauthaler, E. A. B. da Silva, C. L. Pagliari, and S. R. Neves, "Infrared-Visible Image Fusion Using the Undecimated Wavelet Transform with Spectral Factorization and Target Extraction," *submitted for publication in Proceedings of the 2012 IEEE International Conference on Image Processing*.

[4] S. Li, B. Yang, and J. Hu, "Performance comparison of different multi-resolution transforms for image fusion," *Information Fusion*, vol. 12, no. 2, pp. 74–84, 2011.

[5] A. Ellmauthaler, E. A. B. da Silva, and C. L. Pagliari, "Multiscale Image Fusion Using the Undecimated Wavelet Transform With Spectral Factorization and Non-Orthogonal Filter Banks," *submitted for publication in IEEE Transactions on Image Processing*.

[6] D. J. Field, "Scale-invariance and self-similar 'wavelet' transforms: an analysis of natural scenes and mammalian visual systems," in *Wavelets, Fractals and Fourier Transforms: New Developments and New Applications*, pp. 151–193. Oxford University Press, 1993.

[7] A. L. da Cunha, J. Zhou, and M. N. Do, "The nonsubsampled contourlet transform: Theory, design, and applications," *IEEE Transactions on Image Processing*, vol. 15, no. 10, pp. 3089–3101, October 2006.

[8] I. W. Selesnick, R. G. Baraniuk, and N. C. Kingsbury, "The dual-tree complex wavelet transform," *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 123–151, November 2005.

[9] G. Pajares and J. M. de la Cruz, "A wavelet-based image fusion tutorial," *Pattern Recognition*, vol. 37, no. 9, pp. 1855–1872, 2004.

[10] M. J. Shensa, "The discrete wavelet transform: Wedding the à trous and Mallat algorithms," *IEEE Transactions on Signal Processing*, vol. 40, no. 10, pp. 2464–2482, October 1992.

[11] M. Unser, "Ten good reasons for using spline wavelets," in *Proceedings of the 5th SPIE Conference on Wavelet Applications in Signal and Image Processing*, 1997, pp. 422–431.

[12] C. S. Xydeas and V. Petrovic, "Objective image fusion performance measure," *Electronics Letters*, vol. 36, no. 4, pp. 308–309, February 2000.

[13] G. Qu, D. Zhang, and P. Yan, "Information measure for performance of image fusion," *Electronics Letters*, vol. 38, no. 7, pp. 313–315, March 2002.