

MULTIRESOLUTION IMAGE FUSION GUIDED BY A MULTIMODAL SEGMENTATION

¹Gemma Piella, ²Henk Heijmans

¹Gemma.Piella@cwi.nl

¹⁻² Centre for Mathematics and Computer Science (CWI),
Amsterdam, The Netherlands

ABSTRACT

We present a multiresolution fusion algorithm which combines aspects of region and pixel-based fusion. We use multiresolution decompositions to represent the input images at different scales, and introduce a multiresolution/multimodal segmentation to partition the image domain at these scales. This segmentation is then used to guide the subsequent fusion process. A region-based multiresolution approach allows us to consider low-level as well as intermediate-level structures, and to impose data-dependent consistency constraints based on spatial, inter- and intra-scale dependencies.

1. INTRODUCTION

Among the various frameworks in which image fusion has been traditionally formulated, the multiresolution (MR) approach has been one of the most intensively studied and used in practice. The basic idea underlying the MR-based image fusion approach is to perform a MR transform on each source image and, following some specific fusion rules, to construct a composite MR representation from these inputs. The fused image is obtained by applying the inverse transform on this composite MR representation.

Several variants of the MR fusion scheme exist in the literature [1, 2, 3, 4, 5]. These schemes are mainly pixel-based approaches where each individual coefficient of the MR decomposition (or possibly the coefficients in a small fixed neighbourhood) is treated more or less independently. However, for most, if not all, image fusion applications, it could be more meaningful to combine objects rather than pixels. As an intermediate step from pixel-based towards object-based fusion, one might consider region-based approaches. Such approaches have the additional advantage that the fusion process becomes more robust and, moreover, it may help to circumvent some of the well-known drawbacks of pixel-based techniques, such as blurring effects, high sensitivity to noise and misregistration.

In this paper we propose a MR fusion algorithm which combines aspects of region and pixel-level fusion. The basic idea

is to compute a segmentation based on all different source images and to use this segmentation to guide the fusion process. A major difference with other existing region-based approaches [6, 7] is that the segmentation performed is: (1) *multimodal*, in the sense that a single segmentation is obtained based on all source images, and (2) *multiresolution*, in the sense that it is computed in a MR fashion (i.e., it is not merely a segmentation of a sequence of images at different resolutions).

The paper is organized as follows. Section 2 briefly reviews some MR decomposition methods. Section 3 describes a general framework for MR image fusion which supports region-based MR fusion schemes. A particular example is discussed in Section 4. Finally, Section 5 ends with some conclusions.

2. MR DECOMPOSITIONS

A MR scheme decomposes the signal being analyzed into several components, each of which captures information present at a given scale.

2.1. Pyramids

A classical image pyramid [8] consists of a sequence of simplified versions of an original image in which resolution is gradually decreased. In a *low-pass* or *approximation pyramid*, the bottom or zero level $x^{(0)}$ of the pyramid is equal to the original image x , and each successive level image $x^{(k)}$, $k > 0$, is constructed from its lower level by low-pass filtering and subsampling. A *high-pass* or *detail pyramid* can then be derived by interpolation of the approximation images $x^{(k)}$, and subtraction of the outcome from its predecessor $x^{(k-1)}$. At each level, the resulting detail image $y^{(k)}$ contains only those features lost from one level to the next, and therefore contains only details within a restricted range of resolution.

Formally, the *pyramid transform* [9] is the mapping $x^{(0)} \mapsto \{y^{(1)}, y^{(2)}, \dots, y^{(K)}, x^{(K)}\}$, where $y^{(k)}$ is the detail image at level k and $x^{(K)}$ is the approximation at the coarsest level K . Note that the pyramid transform of an image x constitutes a complete representation of x , since the steps used to construct the detail pyramid can be reversed in order to recover

This work is supported by the Dutch Technology Foundation STW, project no. CWI 4616.

x exactly.

There are several ways to modify the pyramid construction method described above. For instance, one can replace the linear filters by morphological ones, which results in a morphological pyramid [10]. Alternatively, one can compute the ratio (instead of the difference) of the successive low-pass filtered images to construct the so-called ratio of low-pass pyramid [11]. Yet another variation on this theme is the gradient pyramid [12], which results when a gradient operator is applied to each level of the low-pass pyramid. In this case, the detail images $y^{(k)}$ are further subdivided into subbands according to their orientation; namely, $y^{(k)} = \{y^{(k)}(\cdot|1), y^{(k)}(\cdot|2), y^{(k)}(\cdot|3), y^{(k)}(\cdot|4)\}$, representing the horizontal, vertical and the two diagonals directions.

2.2. Wavelets

Another MR analysis method similar to pyramids is the *wavelet decomposition* [13]. The main difference is that while pyramids lead to an overcomplete set of transform coefficients, the wavelet decomposition results in a non-redundant representation, i.e., the size of the input equals the size of the output.

One drawback of the (discrete) wavelet transform is that it generally yields a shift-variant signal representation. This means that a simple shift of the input signal may lead to completely different transform coefficients. This lack of translation invariance can be avoided if the outputs of the filter banks are not decimated. The resulting undecimated wavelet transform [13] yields a redundant MR representation where the approximation and detail signals have all the same size as the original input.

Often, wavelet and pyramid transforms are designed for the one-dimensional case. By successive application of such one-dimensional transforms on the rows and the columns (or vice versa) of an image, one obtains a so-called *separable* two-dimensional transform. In the wavelet case, at each level k the approximation image $x^{(k)}$ is decomposed into a coarser approximation $x^{(k+1)}$ and three detail images $y^{(k+1)} = \{y^{(k+1)}(\cdot|1), y^{(k+1)}(\cdot|2), y^{(k+1)}(\cdot|3)\}$, corresponding to the horizontal, vertical and diagonal directions.

Non-separable transforms can also be constructed and provide shift invariance and/or directional selectivity. Moreover, nonlinear extensions of the wavelet transform are possible [14]. A general and flexible approach for the construction of MR decomposition schemes can be found in [9, 15].

2.3. Notation

In this paper, the MR decomposition of an image $x = x^{(0)}$ is denoted by y and it is assumed to be of the form:

$$y = \{y^{(1)}, y^{(2)}, \dots, y^{(K)}, x^{(K)}\}. \quad (1)$$

Here, $x^{(K)}$ represents the approximation image at the highest level (lowest resolution) of the MR structure, while images $y^{(k)}$, $k = 1, \dots, K$, represent the detail images at level k . The detail at level k will, in general, comprise various frequency or orientation bands, depending on the type of MR transform that has been used. We assume henceforth that $y^{(k)}$ is composed of P detail images, i.e., $y^{(k)} = \{y^{(k)}(\cdot|1), \dots, y^{(k)}(\cdot|P)\}$. We use the coordinates (i, j) to index the spatial position of the coefficient. Thus, $x^{(k)}(i, j)$ represents the approximation coefficient at location (i, j) within level k . Similarly, $y^{(k)}(i, j|p)$ represents the detail coefficient at location (i, j) within level k and band p . If no confusion is possible, we will use the shorthand notation of (\cdot) to denote $(i, j|p)$; e.g., we will write $y^{(k)}(\cdot)$ rather than $y^{(k)}(i, j|p)$.

3. A REGION-BASED MR FUSION SCHEME

3.1. Statement of the problem

Before starting the discussion of our region-based fusion scheme, let us briefly recall the problem of image fusion. We are given a set of source images x_S , $S \in \mathcal{S}$, corresponding with a certain scene. These images can come from multiple sensors or from a single sensor device monitored over some period of time. In any case, each of the source images may represent a partial view of the scene, and contain both ‘relevant’ and ‘irrelevant’ data for the task at hand. The goal of an image fusion algorithm is to combine the source images into a single composite image, x_F , which contains a more accurate description of the scene than any of the individual sources. Moreover, the combined image should also appear ‘natural’ so that it can be easily interpreted by humans.

We take as a pre-requisite that the source images are fully registered, and assume we do not have any further information about them than their pixel values, i.e., we do not start from any a priori model. This implies that the localisation of salient information in the source images will depend only on brightness variations.

3.2. MR image fusion

Image fusion requires ‘identification’ of salient information in the sources and its ‘transfer’ into the output (fused) image. The paradigm underlying MR image fusion is that the MR representations of the sources facilitate this type of analysis, not only because it enables one to consider and fuse image features separately at different scales and orientations, but also because it produces large coefficients near edges, thus revealing salient information.

The general structure of a fusion scheme based on MR decompositions is depicted in Fig. 1. First, a MR transform is applied to all source images. Then, a composite MR representation is constructed by a combination of the MR coefficients of the sources. Finally, the fused image is obtained

by computing the inverse transform of the composite representation. As said before, one can find several variants

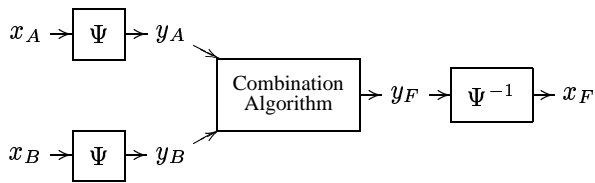


Figure 1: A general MR image fusion scheme comprising a MR transform Ψ of the sources, combination of coefficients in the transform domain, and the inverse multiresolution transform Ψ^{-1} of the composite representation.

of this general scheme in the literature. Most of the existing schemes, however, are pixel-based in the sense that each transform coefficient (or possibly the coefficients within a small fixed window) is considered separately. The main functional blocks of the proposed region-based fusion strategy are depicted in Fig. 2. This scheme is an extension

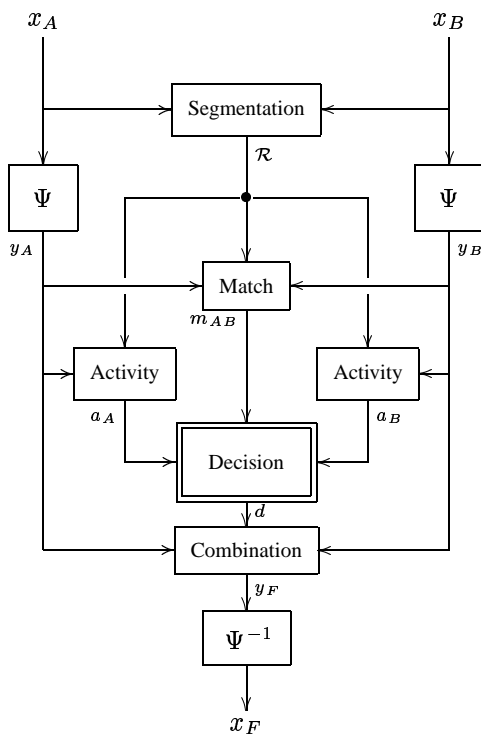


Figure 2: Region-based MR fusion scheme with two input sources x_A and x_B , and one output fused image x_F .

sion of the basic MR fusion scheme of Fig. 1, in which a ‘segmentation block’ has been added and the combination algorithm has been specified. That is, we use MR decompositions to represent the input images at different scales and, addition-

ally, we introduce a multiresolution/multimodal (MR/MM) segmentation to partition the image domain at these scales. In the combination algorithm (see Fig. 1), we distinguish four modules (see Fig. 2): the *activity* and *match* measures extract information from the MR decompositions y_S , which is then used by the *decision* and *combination* maps to compute the MR decomposition y_F of the fused image. The activity measure and match measures are computed for every region in the decomposed input images. These measures may correspond to low-level as well as intermediate-level structures. Furthermore, the MR segmentation \mathcal{R} allows us to impose data-dependent consistency constraints based on spatial as well as inter- and intra-scale dependencies. All this information, i.e, the measures and the consistency constraints, is integrated to yield a decision map d which governs the combination of the coefficients of the transformed sources. This combination results in a MR decomposition y_F , and by MR synthesis we obtain a fused image x_F .

It is interesting to note that for the particular case in which each region corresponds to a single position (i, j) , the region-based approach reduces to a pixel-based approach. Thus, the region-based MR fusion scheme offers a general framework for MR-based image fusion which encompasses most of the existing MR fusion algorithms.

3.3. Description of the modules

In this subsection, we discuss each of the modules in Fig. 2 in more detail. Note, however, that some of them, such as the ‘match module’, are optional. Furthermore, the routines comprised by the various modules can be chosen in a variety of ways.

Multiresolution Analysis

This module computes a MR decomposition of the input sources. Some issues to be addressed at this stage are the specific type of decomposition (pyramid, wavelet, linear, morphological, etc.) and the number of levels. We assume that the same transform Ψ is applied to the source images x_S . Thus, for every input x_S we obtain its MR representation $y_S = \Psi(x_S)$, with y_S having the form defined in (1). For convenience, we will denote the approximation image $x_S^{(K)}$ by $y_S^{(K)}(\cdot|0)$. In this way, we can use the general expression of $y_S^{(k)}(\cdot|p)$ to refer both to the detail images (for $p = 1, \dots, P$) and the approximation image (for $p = 0$).

Multiresolution Multimodal Segmentation

This module uses the various source images as input and returns a single MR segmentation

$$\mathcal{R} = \{\mathcal{R}^{(1)}, \mathcal{R}^{(2)}, \dots, \mathcal{R}^{(K)}\}$$

as output. Here $\mathcal{R}^{(k)}$ represents a segmentation at level k . The basic idea underlying such a segmentation is to identify

the different regions in the scene in order to make a global interpretation of the data at a more abstract level. For this purpose, we develop a MR/MM segmentation algorithm [16] based on a generalized pyramid linking method [17]. Basically, the pyramid linking segmentation algorithm consists of four steps:

1. *Initialization.* Images at lower resolutions are obtained by a filtering/subsampling procedure to construct an approximation pyramid as described in Section 2. We impose the requirement that the sampling scheme used in this stage is the same as the one used in the MR analysis block. However, the approximation pyramid constructed here may be completely different from the MR decomposition obtained in the MR analysis module.
2. *Linking and root labeling.* Child-parent relationships are established between samples at consecutive resolution levels. For each child a suitable parent is sought among the candidate parents: it is linked to its most ‘similar’ parent. We define this similarity based on geometrical and grayscale proximity. Note in particular that ‘similarity’ is computed using all source images. Samples having ‘weak’ parent links and all samples in the highest level are marked as roots.
3. *Segmentation.* The actual segmentation is obtained by using the tree structure of the created links. At each level, all the samples that are connected to a common root are classified as a single region R . Thus, at each level k , we obtain a segmented image $\mathcal{R}^{(k)}$ which contains the different regions R at this level.

Loosely speaking, \mathcal{R} provides a MR representation of the various regions of the underlying scene. This representation will guide the other blocks of the fusion process; hence instead of working at pixel-level, they will take into consideration the regions inferred by the segmentation. As a first approximation, we can regard these regions as the constituent parts of the objects in the overall scene.

Activity Measure

This block associates to every band image $y_S^{(k)}(\cdot|p)$ an activity measure $a_S^{(k)}(\cdot|p)$, which reflects the local activity of the coefficients. Broadly speaking, the activity measure $a_S^{(k)}(i, j|p)$ of a sample (i, j) will be high if the average energy (or some other measure) of $y_S^{(k)}(\cdot|p)$ is high in the vicinity of (i, j) . Thus, one may assume that the activity measure results from some sort of energy calculation over a local neighbourhood of coefficients. In the simplest case, the activity measure is just the absolute value of the coefficient, that is,

$$a_S^{(k)}(\cdot) = |y_S^{(k)}(\cdot)|. \quad (2)$$

Match Measure

The match measure is intended to quantify the degree of ‘similarity’ between the sources. More precisely, the match value $m_{AB}^{(k)}(\cdot)$ reflects the resemblance between the inputs $y_A^{(k)}(\cdot)$ and $y_B^{(k)}(\cdot)$. In the following expression, this value is defined as a normalized correlation averaged over a neighbourhood of the samples:

$$m_{AB}^{(k)}(i, j|p) = \frac{2 \sum_{(u,v) \in \mathcal{W}} w(u, v) y_A^{(k)}(u, v|p) y_B^{(k)}(u, v|p)}{|y_A^{(k)}(i, j|p)|^2 + |y_B^{(k)}(i, j|p)|^2}, \quad (3)$$

where \mathcal{W} is a window centered at (i, j) and w its associated weights.

By analyzing the match measure, one can determine where the sources differ and to which extent, and use this information to combine them in an appropriate way. For example, if the match measure at a given position is low (i.e. the sources are distinctly different at that position), the coefficient from the source decomposition with the highest activity measure is taken as the composite coefficient. On the other hand, if the match measure is high (i.e., the sources are similar at that position), the coefficients from the different sources are averaged to yield the composite coefficient.

Decision Map

The decision map is the main part of the combination algorithm. Its output governs the actual combination of the coefficients of the MR decompositions of the various sources. For each level k , orientation band p , and sample position (i, j) , the decision process assigns a value $d^{(k)}(i, j|p)$ which is then used for the computation of the composite $y_F^{(k)}(i, j|p)$. Most often, the decision takes place independently at each level, band, and position. However, it may also take into account spatial, inter- and intra-scale dependencies between the samples, thus exploiting the idea that coefficients in the composite should not be computed independently. For instance, one may require neighbouring coefficients in the same level and/or orientation to take the same decision.

Combination Map

This module describes the actual combination of the transform coefficients of the sources. For simplicity, let us consider two sources and let us assume that every composite coefficient $y_F^{(k)}(\cdot)$ is ‘assembled’ from the source coefficients at the corresponding level, band and position. More precisely,

$$y_F^{(k)}(\cdot) = C^{(k)} \left(y_A^{(k)}(\cdot), y_B^{(k)}(\cdot), d^{(k)}(\cdot) \right), \quad (4)$$

where $C^{(k)} : \mathbf{R}^3 \mapsto \mathbf{R}$ is the combination map at level k . A simple choice for $C^{(k)}$ is a linear mapping, e.g.,

$$C^{(k)}(y_1, y_2, \delta) = w_A^{(k)}(\delta) y_1 + w_B^{(k)}(\delta) y_2, \quad (5)$$

where the weights $w_A^{(k)}(\delta)$, $w_B^{(k)}(\delta)$ depend on the decision parameter δ .

Henceforth, we assume that the composite coefficients $y_F^{(k)}(\cdot)$ are obtained by a linear combination as in (5), yet with possibly more than two input sources. It remains to be specified how the decision parameter $\delta = d^{(k)}(\cdot)$ determines the weights $w_S^{(k)}(\delta)$. A natural approach is to assign to each coefficient a weight that depends increasingly on the activity measure. In general, the resulting weighted average leads to a stabilization of the fusion result, but it introduces the problem of contrast reduction in case of opposite contrast in different source images. This can be avoided by using a selective combination rule where the most salient component, i.e., the one with largest activity measure, is chosen for the composite. In this case,

$$y_F^{(k)}(\cdot) = y_M^{(k)}(\cdot), \quad (6)$$

with $M = \arg \max_{S \in \mathcal{S}} (a_S^{(k)}(\cdot))$.

In many approaches, the composite approximation coefficients of the highest decomposition level, representing the mean intensity, are taken to be a weighted average of the approximation of the sources, while the composite detail coefficients use a selective combination rule. Additionally, a match measure can be used to determine which combination mode, averaging or selection, to use. Moreover, the decision can be made globally for a group of samples; for example, one might constrain the samples at the same level k and position (i, j) to take the same decision for all bands p . Alternatively, one could 'extend' a decision to a spatial neighbourhood.

Note finally that other factors may be incorporated for the fusion rules. In particular, if some prior knowledge is available, all the fusion blocks can use such information to improve fusion performance. For instance, when combining the source coefficients, the weights assigned to them may depend not only on the activity measure and match measure, but may also reflect some a priori knowledge of an specific type, giving preference to certain levels k , locations (i, j) or input sources.

MR Synthesis

Finally, the fused image is obtained by applying the inverse transformation on the combined MR decomposition y_F :

$$x_F = \Psi^{-1}(y_F), \quad (7)$$

where Ψ^{-1} is the inverse MR transform.

4. A CASE STUDY

Our current implementation of the proposed algorithm is still in an early stage of development. In this section, we consider one of the simplest approaches.

4.1. Specification of the modules

Take two input images x_A and x_B . For their MR decomposition, we use a Laplacian pyramid (thus, we only have a single orientation band, i.e, $P = 1$, and we only distinguish between $p = 0$ and $p = 1$ in this case). In the MR/MM segmentation algorithm, both the linking and root labeling criteria are based exclusively on the difference in grayscale values between parent and child. More sophisticated criteria, as well as connectivity or entropy considerations, are still to be incorporated.

The combination process we consider here does not use a matching measure. We define an activity measure for each region $R \in \mathcal{R}^{(k)}$ in $y_S^{(k)}(\cdot|p)$ by

$$a_S^{(k)}(R|p) = \frac{1}{|R|} \sum_{(i,j) \in R} a_S^{(k)}(i, j|p), \quad (8)$$

where $|R|$ is the size of region R .

The combination process is performed as in (5), where $\delta = d^{(k)}(\cdot)$ and $w_A^{(k)}(\delta) = \delta$, $w_B^{(k)}(\delta) = 1 - \delta$, i.e.,

$$y_F^{(k)}(\cdot) = \delta y_A^{(k)}(\cdot) + (1 - \delta) y_B^{(k)}(\cdot). \quad (9)$$

The decision is computed for each level k and region $R \in \mathcal{R}^{(k)}$ as

$$d^{(k)}(i, j|p) = \begin{cases} 1 & \text{if } a_A^{(k)}(R|p) > a_B^{(k)}(R|p) \\ 0 & \text{otherwise,} \end{cases} \quad (10)$$

for all $(i, j) \in R$.

Thus, the composite images $y_F^{(k)}$ are constructed by a selective combination where the composite coefficients of all samples belonging to a certain region R come from the same source decomposition image (either $y_A^{(k)}$ or $y_B^{(k)}$), namely, the one whose activity in that region is higher. This activity has been computed as the average of the activities measures $a_S^{(k)}(i, j|p)$ with $(i, j) \in R$. For the detail coefficients ($p = 1$) we choose the simple measure $a_S^{(k)}(\cdot) = |y_S^{(k)}(\cdot)|$ (since generally, high magnitude detail coefficient implies relevant perceptual information). For the approximation coefficients ($p = 0$), however, this choice would not be very meaningful as it gives preference to coefficients with high intensity values (bright regions). In the pixel-based schemes found in the literature, one common approach is to disregard the activity measure for the approximation image, and take $\delta = d^{(K)}(i, j|0) = 1/2$ for all (i, j) . This implies that the composite approximation image $y_F^{(K)}(\cdot|0)$ is constructed by the (pixel-wise) average of the approximation images $y_S^{(K)}$. This is also a valid approach for the region-based fusion scheme. However, by doing so, we would be neglecting the region information provided by $\mathcal{R}^{(K)}$. A better choice seems to be the one where the decision map de-

depends on an activity measure based on some texture or entropy criterion. Thus, in our experiments we use the following activity measures:

- For $p = 0$,

$$a_S^{(K)}(R|0) = \frac{1}{|R|} \sum_{(i,j) \in R} \left(y_S^{(K)}(i,j|0) - \bar{y}_S^{(K)}(R|0) \right)^2$$

$$\text{where } \bar{y}_S^{(K)}(R|0) = \frac{1}{|R|} \sum_{(i,j) \in R} y_S^{(K)}(i,j|0).$$

- For $p = 1$,

$$a_S^{(k)}(R|1) = \frac{1}{|R|} \sum_{(i,j) \in R} |y_S^{(k)}(i,j|1)|.$$

We have tested our algorithm on several pairs of images. Two examples are given here to illustrate the fusion process described above. In both cases, we have chosen $K = 3$ and, when displaying the images, the gray values of the pixels have been scaled between 0 and 255 (histogram stretching). The input sources x_A and x_B are displayed, respectively, on the left and right top of the corresponding figure. For the decision maps, pixels with $\delta = 0, 1$ are displayed in black and white, respectively. Thus, according to our algorithm, coefficients corresponding to ‘white regions’ are selected from A , while coefficients corresponding to ‘black regions’ are selected from B .

4.2. First experiment

Fig. 3 shows the fusion of images obtained with different sensors. Note that in the visual image (Fig. 3, top left) it is hard to distinguish the person in camouflage from the background, while it is clearly observable in the infrared (IR) image (Fig. 3, top right). In contrast, the easily discernible background in the visual image is nearly undetectable in the IR image. The first level of the resulting segmentation and decision map are shown in the middle row. The corresponding second and third levels are displayed on the bottom left. It is interesting to note that, according to $d^{(2)}$ and $d^{(3)}$, although most of the background is selected from their respective visual image $y_A^{(k)}$, $k = 2, 3$, the region corresponding to the person is selected from their respective IR image $y_B^{(k)}$, $k = 2, 3$. The final fused image is depicted at the bottom right of Fig. 3.

4.3. Second experiment

The example in Fig. 4 illustrates the usage of fusion in radiotherapy and skull surgery. Here, the information provided by magnetic resonance imaging (MRI) and X-ray computed

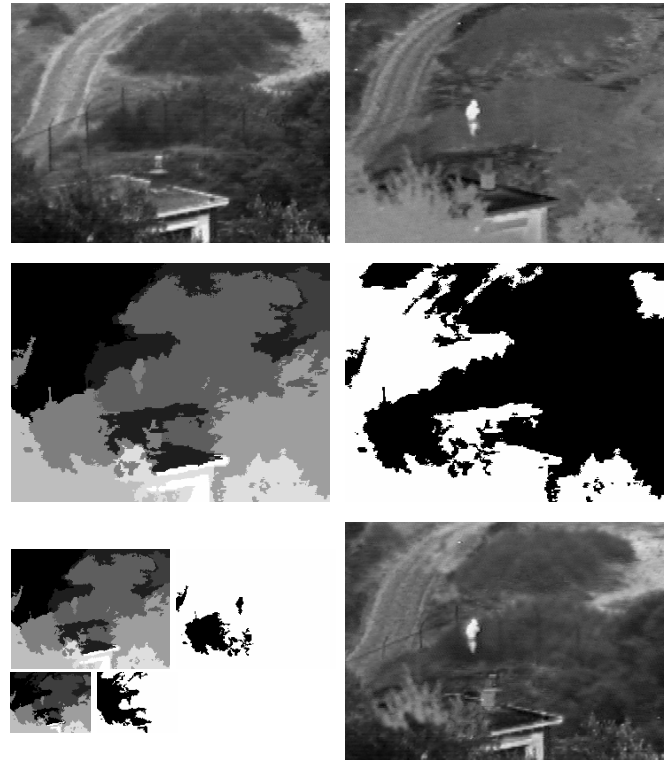


Figure 3: *First experiment. Top: visual and infrared source images. Middle: 1st level of segmentation (left) and decision map (right). Bottom: 2nd and 3rd level of segmentation (left) and decision map (middle), and fused image (right). Source images are courtesy of Alexander Toet, TNO Human Factors Institute, The Netherlands*

tomography (CT) is complementary. Normal and pathological soft tissues are better visualized by MRI (Fig. 4, top left), while the structure of the tissue bone is better visualized by CT (Fig. 4, top right). The fused image, depicted at the bottom right of Fig. 4, provides salient information from both images simultaneously, which may be useful for physicians in medical diagnosis. For this particular example, we also show (Fig. 4, bottom left) the corresponding fused output we would have obtained using a pixel-based MR fusion algorithm with the same fusion rules (except for $p = 0$ where we use a fix $\delta = 1/2$) as in the region-based algorithm.

4.4. Discussion

Despite the crudeness of the current implementation, the visual performance is surprisingly good. This suggests that the region-based approach proposed here can at least be competitive with (but more likely outperform) other MR fusion techniques.

Further investigations are necessary for the fine-tuning of parameters as well as the proper selection of the different ingredients of the scheme. Towards this end, performance assess-

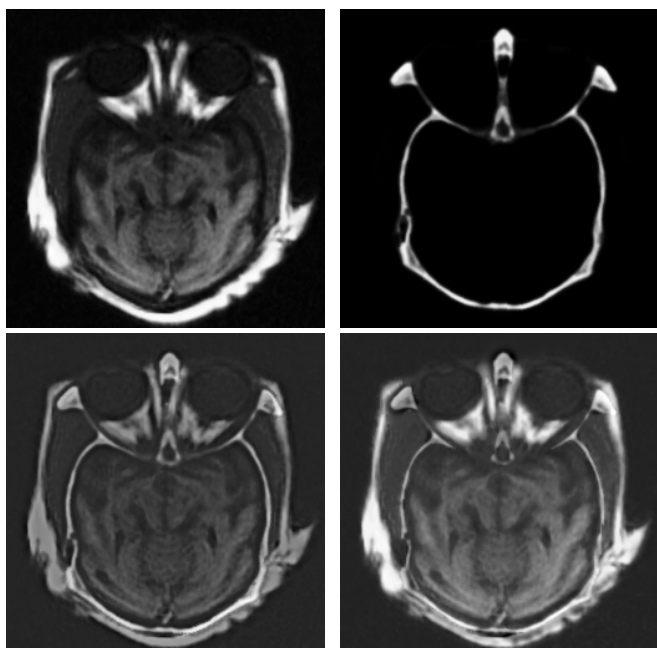


Figure 4: Second experiment. Top: MRI and CT source images. Bottom: pixel-based (left) and region-based (right) fused images.

ment criteria will be developed to evaluate and demonstrate the capacities of the new fusion technique, as well as to compare its performance with others MR fusion schemes. Tests will be carried out based both on objective and subjective criteria.

5. CONCLUSIONS

In this paper, we have introduced a general framework for MR image fusion which supports both pixel and region-based approaches. The implementation of our algorithm is still in a preliminary stage and in the experiments performed we did not attempt to optimize its performance. However, the results obtained so far suggest that our approach may be useful for several image fusion applications. We intend to investigate this more thoroughly in the future. A substantial part of our efforts will be devoted to the design of objective measures for fusion performance assessment.

In the future we also intend to replace the MR/MM segmentation by pyramid linking by some other approach, such as a hierarchical watershed algorithm from mathematical morphology.

6. REFERENCES

- [1] P. J. Burt and R. J. Kolczynski, "Enhanced image capture through fusion," in *Proceedings of the 4th International Conference on Computer Vision*, Berlin, Germany, May 1993, pp. 173–182.
- [2] H. Li, B. S. Manjunath, and S. K. Mitra, "Multisensor image fusion using the wavelet transform," *Graphical Models and Image Processing*, vol. 57, no. 3, pp. 235–245, May 1995.
- [3] I. Koren, A. Laine, and F. Taylor, "Image fusion using steerable dyadic wavelet transforms," in *Proceedings of the IEEE International Conference on Image Processing*, Washington D.C., October 1995, pp. 232–235.
- [4] T. A. Wilson, S. K. Rogers, and L. R. Meyers, "Perceptual based hyperspectral image fusion using multiresolution analysis," *Optical Engineering*, vol. 34, no. 11, pp. 3154–3164, 1995.
- [5] L. J. Chipman and T. M. Orr, "Wavelets and image fusion," in *Proceedings of the IEEE International Conference on Image Processing*, Washington D.C., October 1995, pp. 248–251.
- [6] B. J. Matuszewski, L.-K. Shark, M. R. Varley, and J. P. Smith, "Region-based wavelet fusion of ultrasonic, radiographic and shearographic non-destructive testing images," in *Proceedings of the 15th World Conference on Non-Destructive Testing*, Rome, October 2000.
- [7] Z. Zhang and R. Blum, "A region-based image fusion scheme for concealed weapon detection," in *Proceedings of the 31th Annual Conference on Information Sciences and Systems*, March 1997, pp. 168–173.
- [8] P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Transactions on Communications*, vol. 31, pp. 532–540, 1983.
- [9] J. Goutsias and H. J. A. M. Heijmans, "Nonlinear multiresolution signal decomposition schemes. Part I: Morphological pyramids," *IEEE Transactions on Image Processing*, vol. 9, no. 11, 2000.
- [10] A. Toet, "A morphological pyramidal image decomposition," *Pattern Recognition Letters*, vol. 9, pp. 255–261, 1989.
- [11] A. Toet, "Image fusion by a ratio of low-pass pyramid," *Pattern Recognition*, vol. 9, pp. 245–253, 1989.
- [12] P. J. Burt, "A gradient pyramid basis for pattern selective image fusion," in *Proceedings of the Society for Information Display Conference*, 1992.
- [13] S. G. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, California, 1998.
- [14] R. L. de Queiroz, D. A. F. Florêncio, and R. W. Schafer, "Nonexpansive pyramid for image coding using a nonlinear filterbank," *IEEE Transactions on Image Processing*, vol. 7, pp. 246–252, 1998.

- [15] H. J. A. M. Heijmans and J. Goutsias, “Nonlinear multiresolution signal decomposition schemes. Part II: Morphological wavelets,” *IEEE Transactions on Image Processing*, vol. 9, no. 11, pp. 1897–1913, 2000.
- [16] G. Piella, “A general framework for multiresolution image fusion: from pixels to regions,” Research Report PNA-R0211, CWI, Amsterdam, 2002.
- [17] P. J. Burt, T. H. Hong, and A. Rosenfeld, “Segmentation and estimation of image region properties through cooperative hierarchical computation,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 11, no. 12, pp. 802–809, December 1981.