A simple and efficient algorithm for multifocus image fusion using morphological wavelets

Ishita De[a], Bhabatosh Chanda[b,*]

[a]Department of Computer Science, Barrackpore Rastraguru Surendranath College, Kolkata 700120, India
[b]Electronics and Communication Sciences Unit, Indian Statistical Institute, Kolkata 700108, India

Abstract

This paper presents a simple yet efficient algorithm for multifocus image fusion, using a multiresolution signal decomposition scheme. The decomposition scheme is based on a nonlinear wavelet constructed with morphological operations. The analysis operators are constructed by morphological dilation combined with quadratic downsampling and the synthesis operators are constructed by morphological erosion combined with quadratic upsampling. A performance measure based on image gradients is used to evaluate the results. The proposed scheme has some interesting computational advantages as well.

Keywords: Multifocus image fusion; Multiresolution signal decomposition; Morphological wavelets; Image gradient

1. Introduction

A wide variety of data acquisition devices are available at present, and hence image fusion has become an important subarea of image processing. There are sensors which cannot generate images of all objects at various distances (from the sensor) with equal clarity (e.g. camera with finite depth of field, light optical microscope, etc.). Thus several images of a scene are captured, with focus on different parts of it. The acquired images are complementary in many ways and a single one of them is not sufficient in terms of their respective information content. However, viewing a series of such images separately and individually is not very useful and convenient. The advantages of *multi-focus data* can be fully exploited by integrating the sharply focused regions seen in the different images. Before integrating, one must bring the constituent images to a common coordinate system. Widely used techniques perform this task by using some common geometrical references called ground control points (GCPs). This process is called registration. After registration the images

are combined to form a single image through a judicious selection of regions from different images. This process is known as *multifocus image fusion*. Both the reliability of redundant information and the quality of complementary information present in the constituent images are improved in a fused image. So it gives a better view for human and/or machine perception. A fused data can also render itself more successfully for any subsequent processing like object recognition, feature extraction, segmentation, etc.

There are a number of techniques for multifocus image fusion. Simple techniques in which the fusion operation is performed directly on the source images (e.g. weighted average method), often have serious side effects like reduction in the contrast of the fused image. Other approaches include, image fusion using controllable camera [1], probabilistic methods [2], image gradient method with majority filtering [3], multiscale methods [4] and multiresolution approaches [5–8]. Methods described in [1] depend on controlled camera motion and do not work for arbitrary set of images. Probabilistic techniques involve huge computation using floating point arithmetic and thus requires a lot of time and memory-space. Image gradient method with majority filtering has the drawback that the defocused zone of one image is enhanced at the expense of focused zone of others.

An image often contains physically relevant features at many different scales or resolutions. Multiscale and multiresolution approaches provide a means to exploit this fact. This is one of the reasons why these techniques have become so popular. Multiscale methods involve processing and storing of scaled data at various levels which are of the same size as that of the original images. This results in a huge amount of memory and time requirement [4]. Multiresolution techniques of image fusion using pyramid or wavelet transform produce good result in less computation time using less memory. The basic idea in these techniques is to decompose the source images at first by applying the pyramid or wavelet transform, then the fusion operation on the transformed images is performed and finally the fused image is reconstructed by inverse transform. One major advantage of multiresolution transform is that *spatial* as well as *frequency* domain localization of an image is obtained simultaneously. Another advantage is that it can provide information on the sharp contrast changes, and human visual system is especially sensitive to these changes. Both pyramid and wavelet transforms are used as multiresolution filters. Wavelet transform can be considered as a special case of pyramid transform but it has more complete theoretical support [9]. Various methods of image fusion using multiresolution techniques have been suggested before [5–8]. A survey on these works may be found in [8]. Burt and Lolezynski [5] suggested a method in which the images are first decomposed into a gradient pyramid. Activity measure of each pixel is computed then by finding out the variance of a 3×3 or 5×5 window centered around that pixel. Depending on this measure, either the larger value or the average value is chosen. Finally the reconstruction is done. Li et al. [6] used similar method except the fact that wavelet transforms are used for decomposition and consistency verification is done along with area-based activity measure and maximum selection. Their method reduces the artifacts such as blocking effects which are common in image fusion using multiresolution. Methods described in [5,6] are complex and time-consuming. Moreover it was not mentioned whether the method can be applied to more than two multifocus images. In the method suggested by Yang et al. [7], an impulse function is defined at first to describe the quality of an object in a multifocus image. Then sharply focussed regions are extracted by analyzing the wavelet coefficients of two primary images and two blurred images. To fuse two images, this method compares the wavelet coefficients of four images and thus involves double computation. Moreover, all its computations involve floating point arithmetic.

Wavelet transform is a linear tool in its original form [9]. But *nonlinear* extensions of discrete wavelet transform are possible by various methods like *lifting scheme* [10] or *morphological operators* [11,12]. The problem with linear wavelets like Haar wavelet is that during signal decomposition or analysis the range of the original data is not preserved [12]. Secondly, linear wavelets act as

low-pass filters and thus smooth-out the edges. This results in reduction in the contrast in fused images. The nonlinear wavelet introduced by Heijmans et al. [12] overcomes this drawback by using morphological operators. But it involves division operation and thus either requires floating point arithmetic or introduces truncation error by using integer arithmetic.

We introduce in this paper a *nonlinear morphological wavelet transform* which preserves the range in the scaled images and involves integer arithmetic only. We use this transform to present a fusion algorithm to fuse a set of two-dimensional gray-scale multifocus images of some scene. The method is simple, computationally efficient and produces good results. Integrated-chip implementations of image processing algorithms are going to become more common in near future. Our method will be useful in this respect. The results obtained by it has been compared with those obtained by using Haar wavelet and the morphological wavelet suggested by Heijmans and Goutsias [12] using several sets of data. The paper is organized as follows. Section 2 gives the basic theory (without proof) of multiresolution analysis using wavelets and a brief discussion on morphological operators. This section also introduces the proposed wavelet transform based on these operators. Section 3 describes the image-fusion algorithm using the new morphological wavelet. Experimental results and discussion are given in Section 4 and the concluding remarks are presented in Section 5.

2. Basic theory and a new morphological wavelet

A brief overview of multiresolution signal decomposition theory using wavelets is given first, followed by the discussion on morphological operators, and finally a new wavelet transform based on morphological operators is presented.

2.1. Multiresolution analysis

The theory of multiresolution signal decomposition scheme using wavelets can be applied to a wide variety of signals. We are restricted here to two-dimensional gray-scale image signals only. A two-dimensional gray-scale image signal X is a function which map (a subset of) discrete two-dimensional space Z^2 to a finite set of non-negative integers $G = \{g_1, g_2, \ldots, g_n\}$ called the set of gray values. Let us consider a set V_0 of such image signals. A multiresolution signal decomposition scheme on V_0 uses two types of operators, namely, *signal analysis* and *signal synthesis* operators. *Signal analysis* operators $\psi_j^{\uparrow} : V_j \to V_{j+1}$, map the signal space V_j at level j, to a coarser signal space V_{j+1} and the *detail analysis* operators $\omega_j^{\uparrow} : V_j \to W_{j+1}$, map V_j to a coarser detail space W_{j+1}. All V_j's and W_j's have the same structure as V_0. The operators $\psi_j^{\uparrow}$ and $\omega_j^{\uparrow}$ are called the *scaling function* and the *wavelet function*, respectively. Signal analysis operation proceeds by mapping a signal to a level higher in the pyramid structure, thereby reducing information. Details are stored at each level to restore this information loss. So, for $X^j \in V_j$ we have,

$$\psi_j^{\uparrow}(X^j) = X^{j+1}, \quad X^{j+1} \in V_{j+1}, \tag{1}$$

$$\omega_j^{\uparrow}(X^j) = Y^{j+1}, \quad Y^{j+1} \in W_{j+1}. \tag{2}$$

Signal synthesis or reconstruction is done by *synthesis operator* $\Psi_j^{\downarrow} : V_{j+1} \times W_{j+1} \to V_j$, which map a signal to a level lower in the pyramid. To ensure *loss-less* or *perfect reconstruction*, the following condition must be satisfied.

$$\Psi_j^{\downarrow}(\psi_j^{\uparrow}(X^j), \omega_j^{\uparrow}(X^j)) = X^j, \quad X^j \in V_j. \tag{3}$$

There are two more conditions, namely,

$$\psi_j^{\uparrow}(\Psi_j^{\downarrow}(X^{j+1}, Y^{j+1})) = X^{j+1}, \tag{4}$$

$$\omega_j^{\uparrow}(\Psi_j^{\downarrow}(X^{j+1}, Y^{j+1})) = Y^{j+1}, \tag{5}$$

where $X^{j+1} \in V_{j+1}$ and $Y^{j+1} \in W_{j+1}$. They ensure that the decomposition is *non-redundant* in the sense that repeated applications of these schemes produce the same result. A special case called *uncoupled* wavelet decomposition occurs when there exists a binary operation $\dot{+}$ on V_j and operators $\psi_j^{\downarrow} : V_{j+1} \to V_j$ and $\omega_j^{\downarrow} : W_{j+1} \to V_j$

such that

$$\Psi_j^{\downarrow}(X^{j+1}, Y^{j+1}) = \psi_j^{\downarrow}(X^{j+1}) \dot{+} \omega_j^{\downarrow}(Y^{j+1}),$$
$$X^{j+1} \in V_{j+1}, \quad Y^{j+1} \in W_{j+1}. \tag{6}$$

Then perfect reconstruction and non-redundancy conditions become

$$\psi_j^{\downarrow}\psi_j^{\uparrow}(X^j) \dot{+} \omega_j^{\downarrow}\omega_j^{\uparrow}(X^j) = X^j, \quad X^j \in V_j, \tag{7}$$

$$\psi_j^{\uparrow}(\psi_j^{\downarrow}(X^{j+1}) \dot{+} \omega_j^{\downarrow}(Y^{j+1})) = X^{j+1},$$
$$X^{j+1} \in V_{j+1}, \quad Y^{j+1} \in W_{j+1}, \tag{8}$$

$$\omega_j^{\uparrow}(\psi_j^{\downarrow}(X^{j+1}) \dot{+} \omega_j^{\downarrow}(Y^{j+1})) = Y^{j+1},$$
$$X^{j+1} \in V_{j+1}, \quad Y^{j+1} \in W_{j+1}. \tag{9}$$

If an one-dimensional wavelet decomposition scheme can be applied to two and higher dimensions, by applying it to other dimensions sequentially, then this decomposition is called *separable*.

2.2. Morphological operators

A brief overview of the morphological operators is given now. Let us consider a signal $X \in V_0$ (see Section 2.1). So X is a function from domain D to G where D is a subset of Z^2 and G is the set of gray-values. Let $A \subset Z^2$ be a *structuring element*. Then the morphological operators, *dilation* $\delta_A(X)$ and *erosion* $\varepsilon_A(X)$ of X by A are defined as

$$\delta_A(X)(r,c) = \max_{(r_1,c_1)\in A,(r-r_1,c-c_1)\in D} X(r-r_1, c-c_1),$$

$$\varepsilon_A(X)(r,c) = \min_{(r_1,c_1)\in A,(r+r_1,c+c_1)\in D} X(r+r_1, c+c_1).$$

So *dilation* (*erosion*) simply replace the value at each point of X by the maximum (minimum) value in a neighborhood defined by the *structuring element* A. Other morphological operators can be constructed by combining δ and ε. For example *opening* $\alpha_A(X)$ and *closing* $\phi_A(X)$ of X by A are defined as

$$\alpha_A(X)(r,c) = \delta_A(\varepsilon_A(X))(r,c),$$

$$\phi_A(X)(r,c) = \varepsilon_A(\delta_A(X))(r,c).$$

2.3. A new morphological wavelet

Heijmans and Goutsias introduced a *morphological variant* of the linear Haar wavelet by using the morphological operation *dilation* (*erosion*) [12]. It is an one-dimensional scheme and the multi-dimensional implementation can be obtained by applying it to other dimensions sequentially. However, a non-separable two-dimensional version of the morphological Haar wavelet transform has also been defined in [12], which will be used in our experiments. We, now propose a *non-separable two-dimensional uncoupled morphological wavelet* decomposition scheme, which will be used for our image-fusion algorithm. Unique analysis operators $(\psi^{\uparrow}, \omega^{\uparrow})$ are used at all levels of the multiresolution scheme. Similarly, unique synthesis operators $(\psi^{\downarrow}, \omega^{\downarrow})$ are used at all levels. These operators are explained for the lowermost levels 0 and 1.

Let us consider the signal space V_0 of Section 2.1. It is our original signal space. Then V_1 and W_1 are the signal and detail spaces at level 1 having the same structure as V_0. Consider an image signal $X \in V_0$. Then X is a mapping of (a subset of) Z^2 to the set of gray-values G and it can be represented by an $M \times N$ matrix, where $M, N \in Z$. Let us assume that M and N both are even. Then X can be divided into consecutive and disjoint 2×2 submatrices or blocks, which are total $MN/4$ in number. Four positions of such a block B may be denoted by (r,c), $(r,c+1)$, $(r+1,c)$ and $(r+1,c+1)$ (see Fig. 1). Using quadratic downsampling, the analysis operators $\psi^{\uparrow} : V_0 \to V_1$ and $\omega^{\uparrow} : V_0 \to W_1$ are defined as

$$\psi^{\uparrow}(X)(B) = M = \max\{X(r,c), X(r,c+1), X(r+1,c), X(r+1,c+1)\}, \tag{10}$$

$$\omega^{\uparrow}(X)(B) = (y_v, y_h, y_d), \tag{11}$$

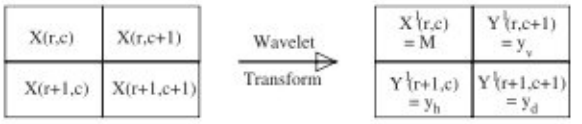


Fig. 1. Wavelet transform on a 2×2 block of X.

where y_v, y_h, y_d represent the *vertical*, *horizontal* and *diagonal* detail signals, respectively, defined as

$$y_v = \begin{cases} M - X(r, c+1) & \text{if } M - X(r, c+1) > 0, \\ X(r, c+1) - M & \text{otherwise,} \end{cases} \tag{12}$$

$$y_h = \begin{cases} M - X(r+1, c) & \text{if } M - X(r+1, c) > 0, \\ X(r+1, c) - M & \text{otherwise,} \end{cases} \tag{13}$$

$$y_d = \begin{cases} M - X(r+1, c+1) & \text{if } M - X(r+1, c+1) > 0, \\ X(r+1, c+1) - M & \text{otherwise.} \end{cases} \tag{14}$$

The second condition in the last three equations is required to maintain the information on position of the maximum value M as evidenced in the successive example. Scaled signal and detail values obtained above belong to X^1 and Y^1, respectively, and they can be stored conveniently in similar positions of another matrix.

The original signal at level 0 is reconstructed by the synthesis operation. Using quadratic down-sampling, *synthesized signals* $\widehat{X}$ are given by

$$\begin{aligned} \widehat{X}(r, c) &= \widehat{X}(r, c+1) = \widehat{X}(r+1, c) \\ &= \widehat{X}(r+1, c+1) = M \end{aligned} \tag{15}$$

and *synthesized details* $\widehat{Y}$ are given by

$$\widehat{Y}(r, c) = \min(y_v, y_h, y_d, 0), \tag{16}$$

$$\widehat{Y}(r, c+1) = \min(-y_v, 0), \tag{17}$$

$$\widehat{Y}(r+1, c) = \min(-y_h, 0), \tag{18}$$

$$\widehat{Y}(r+1, c+1) = \min(-y_d, 0), \tag{19}$$

where $M = X^1(r, c)$ is the scaled signal at (r, c) and y_v, y_h, y_d are vertical, horizontal and diagonal details, respectively. This is an uncoupled decomposition scheme and the binary operation $\dot{+}$ is the ordinary addition of numbers. Hence the reconstructed signal X' at any point $(u, v) \in \{(r, c), (r, c+1), (r+1, c), (r+1, c+1)\}$ is given by

$$X'(u, v) = \widehat{X}(u, v) \dot{+} \widehat{Y}(u, v). \tag{20}$$

Example. Let us consider the 2×2 block B of X with $X(r, c) = T_0, X(r, c+1) = T_1, X(r+1, c) = T_2$ and $X(r+1, c+1) = T_3$. Let $T_m = \max\{T_0, T_1, T_2, T_3\}$. Then $\psi^{\uparrow}(X)(B) = T_m$ and the details are given by $\omega^{\uparrow}(X)(B) = (T_v, T_h, T_d)$ where

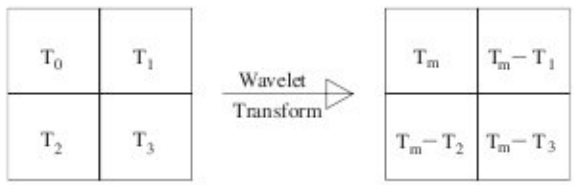
$$T_v = \begin{cases} T_m - T_1 & \text{if } T_m - T_1 > 0, \\ T_0 - T_m & \text{otherwise,} \end{cases}$$

$$T_h = \begin{cases} T_m - T_2 & \text{if } T_m - T_2 > 0, \\ T_0 - T_m & \text{otherwise,} \end{cases}$$

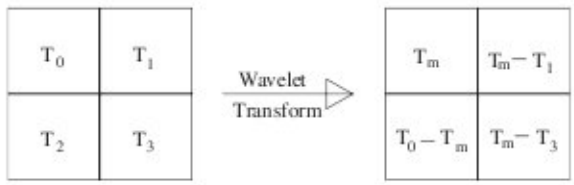
$$T_d = \begin{cases} T_m - T_3 & \text{if } T_m - T_3 > 0, \\ T_0 - T_m & \text{otherwise.} \end{cases}$$

Now T_m may occur at any of the four positions of the block 2×2 submatrix. The situations of T_m occurring at (r, c) and $(r+1, c)$ are illustrated in the Fig. 2. In the first case T_m occurs at position (r, c) and all the detail values are positive. In the second case T_m occurs at position $(r+1, c)$ and the information is preserved by placing the negative value $T_0 - T_m$ as the horizontal detail.

The analysis operator-pair $(\psi_j^{\uparrow}, \omega_j^{\uparrow})$ can be used recursively to decompose a signal upto a desired level $k \geqslant 1$. Similarly the synthesis operator-pair $(\psi_j^{\downarrow}, \omega_j^{\downarrow})$ can be used recursively to reconstruct a signal from any level to the lowest level 0. It is easy to see that the analysis and synthesis operators



Case 1: Transform when T_0 is maximum



Case 2: Transform when T_2 is maximum

Fig. 2. Example for proposed wavelet transform.

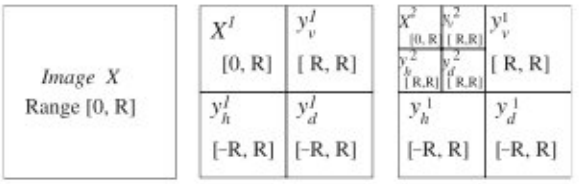


Fig. 3. (a) Original signal X, (b) scaled signal X^1 and details $Y^1 = \{y_v^1, y_h^1, y_d^1\}$ at level 1, (c) scaled signal X^2 and details $Y^2 = \{y_v^2, y_h^2, y_d^2\}$ at level 2.

satisfy the perfect reconstruction and non-redundancy conditions 7–9 given in Section 2.1. The operators $\psi^{\uparrow}$ and $\omega^{\uparrow}$ involve simple arithmetic operations and one interesting point to note is that the integer values are mapped to integer values only. Another point to note is that, if all values of X belong to the range $[0, R]$, then analyzed signal-values will belong to the range $[0, R]$ and analyzed detail-values will belong to the range $[-R, R]$, irrespective of the number of times the operators are applied (see Fig. 3).

3. Image fusion

We now present the image fusion algorithm proposed by us using the morphological wavelet transform given in Section 2.3. Consider n two-dimensional multifocus images $X_1, X_2, \ldots, X_n$. These images must be registered and of same size. The proposed analysis operators $\psi^{\uparrow}$ and $\omega^{\uparrow}$, are applied on the n individual images k times recursively. If $X_i, i = 1, 2, \ldots, k$ are $M \times N$ images, the analysis operators can be applied at most $k_{\max}$ times where $k_{\max} = \min(\lfloor \log_2 M \rfloor, \lfloor \log_2 N \rfloor)$. After completion of the analysis operation, a set of *n scaled image signals* at the topmost level k are obtained. They are denoted by $X_i^k, i = 1, 2, \ldots, n$. A set of *detail signals* Y_i^j, $i = 1, 2, \ldots, n$ are also obtained at each level $j, j = 1$–k. As mentioned in the last section, if the range of image X_i is $[0, R]$, then the range of scaled image X_i^k is $[0, R]$ and the range of details Y_i^j, $j = 1, 2, \ldots, k$ are $[-R, R]$. While comparing $X_i^k, i = 1, 2, \ldots, n$ positionwise, greater absolute value corresponds to a brighter pixel and while comparing Y_i^j, $j = 1, 2, \ldots, k, i = 1, 2, \ldots, n$ positionwise, greater absolute value corresponds to sharp contrast changes such as edge, line and region boundaries. Based on this observation, scaled images $X_i^k, i = 1, 2, \ldots, n$ are combined by comparing the values at each position (r, c) and choosing the one with greatest absolute value. Similar operation is applied on corresponding details at each level. Thus a single fused image at level k and a detail at each level $j, j = 1, \ldots, k$ are obtained. Then the reconstruction phase begins. The image at level $k-1$ is reconstructed by applying the synthesis operators $\psi^{\downarrow}$ and $\omega^{\downarrow}$ (as proposed by us in the previous section) followed by addition. Synthesis operators are applied k times recursively to obtain the image at level 0. The algorithm can be summarized as below.

3.1. The algorithm

1. *Analysis step*: Apply the analysis operators $\psi^{\uparrow}$ and $\omega^{\uparrow}, k$ times recursively, on image $X_i, i = 1, \ldots, n$ and get $\mathbf{X_i} = \{X_i^k, Y_i^1, Y_i^2, \ldots, Y_i^k\}$, where X_i^k is the scaled image at level k and Y_i^j, $j = 1, \ldots, k$ are the details at levels $1, 2, \ldots, k$, respectively.
2. *Fusion step*: Compare $\{\mathbf{X_i}, i = 1, 2, \ldots, n\}$ and combine them into $\mathbf{X} = \{X^k, Y^1, Y^2, \ldots, Y^k\}$, where X^k and Y^j are given by $X^k(r, c) = \max\{|X_1^k(r, c)|, |X_2^k(r, c)|, \ldots, |X_n^k(r, c)|\}$ and $Y^j(r, c) = \max\{|Y_1^j(r, c)|, |Y_2^j(r, c)|, \ldots, |Y_n^j(r, c)|\}$, respectively.
3. *Synthesis step*: Reconstruct the fused image X^j at level j, $j = k-1, \ldots, 0$, by applying the synthesis operators $\psi^{\downarrow}$ and $\omega^{\downarrow}$, respectively, on X^{j+1} and Y^{j+1} following by addition, i.e.

$$X^j(r, c) = \psi^{\downarrow}(X^{j+1}(r, c)) + \omega^{\downarrow}(Y^{j+1}(r, c)).$$

3.2. Example

The algorithm is illustrated by using 2×2 sample data A and B taken from the multifocus images X_1 and X_2, respectively.

Let

$$A = \begin{bmatrix} a_0 & a_1 \\ a_2 & a_3 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} b_0 & b_1 \\ b_2 & b_3 \end{bmatrix},$$

where a_i and $b_i, i = 0,1,2,3$ are non-negative integers. Applying the analysis operators $\psi^{\uparrow}$ and $\omega^{\uparrow}$ once, A becomes,

$$A^1 = \begin{bmatrix} a_0^1 & a_1^1 \\ a_2^1 & a_3^1 \end{bmatrix},$$

where

$a_0^1 = a_{\max} = \max\{a_i, i = 0,1,2,3\}$ and

$$a_i^1 = \begin{cases} a_{\max} - a_i & \text{if } a_{\max} > a_i, \\ -(a_{\max} - a_0) & \text{otherwise.} \end{cases}$$

Here a_0^1 is the scaled signal-data and $a_i^1, i = 1,2,3$ are the detail-data at level 1.

Similarly, after the analysis operation, B becomes,

$$B^1 = \begin{bmatrix} b_0^1 & b_1^1 \\ b_2^1 & b_3^1 \end{bmatrix}$$

A^1 and B^1 are fused in C^1, by the fusion step, where

$$C^1 = \begin{bmatrix} c_0^1 & c_1^1 \\ c_2^1 & c_3^1 \end{bmatrix} \quad \text{and}$$

$$c_i^1 = \begin{cases} a_i^1 & \text{if } |a_i^1| \geqslant |b_i^1|, \\ b_i^1 & \text{otherwise,} \end{cases} \quad \text{for } i = 0,1,2,3.$$

The fused data C at level 0 is obtained by applying the synthesis operators $\psi^{\downarrow}$ and $\omega^{\downarrow}$ followed by addition. Therefore

$$C = \begin{bmatrix} c_0 & c_1 \\ c_2 & c_3 \end{bmatrix},$$

where

$$c_0 = c_0^1 + \min(0, c_1^1, c_2^1, c_3^1) \quad \text{and}$$

$$c_i = \begin{cases} c_0^1 & \text{if } c_i^1 < 0, \\ c_0^1 - c_i^1 & \text{otherwise,} \end{cases} \quad \text{for } i = 1,2,3.$$

We now, claim that, c_i is always less or equal to R, where R is the greatest value of a_i and $b_i, i = 0,1,2,3$. This happens because c_i is obtained by subtracting a non-negative value from $c_0^1 = \max(a_0^1, b_0^1)$. However the lower bound of c_i may not remain within the lower bounds of A and B. The method can be applied to the complete images X_1 and X_2 by taking as many 2×2 samples as required.

4. Experimental results and discussion

The proposed method for image fusion has been implemented in C language on Unix environment and has been tested on a number of multifocus images. We have compared our method with similar methods using Haar wavelet and the two-dimensional morphological Haar wavelet introduced by Heijmans and Goutsias [12]. The experimental results are shown in Figs. 4–7. In each figure, the original multifocus images are given first, followed by the fused images obtained by the proposed method and two other methods mentioned above. The fusion is done by decomposing the constituent images up to the third level, in all the cases. Time taken by these algorithms are more or less same and not significant.

4.1. Performance analysis

Careful manual inspection of Figs. 4–7 reveals that the results obtained by the proposed wavelet are better than that of Haar wavelet and are comparable to that of Heijmans and Goutsias' method [12]. However this is a subjective measure of quality and may not be universally acceptable. Hence a quantitative measure is also given.

Gradient or derivative operators are useful tools to measure the variation of intensity with respect to immediate neighboring points or pixels of an image [13]. It is observed that a pixel possesses high gradient value when it is sharply focused. An objective criterion based on this knowledge is suggested to measure the quality of the results. The gradient operator suggested by Roberts [14] is used here. Thus magnitude of gradient $G(r,c)$ at a point (r,c) of image X is obtained by

$$G(r,c) = \tfrac{1}{2}\{|X(r,c) - X(r+1,c+1)| + |X(r,c+1) - X(r+1,c)|\}.$$

It may be noted that $G(r,c)$ is obtained from X through a non-linear transformation. So, we

Fig. 4. Original multifocus images and the fused images by different algorithms: (a) Near focused image; (b) middle focused image; (c) far focused image; (d) fused image with proposed algorithm; (e) fused image by Haar wavelet; (f) fused image by Heijmans and Goutsias' method.

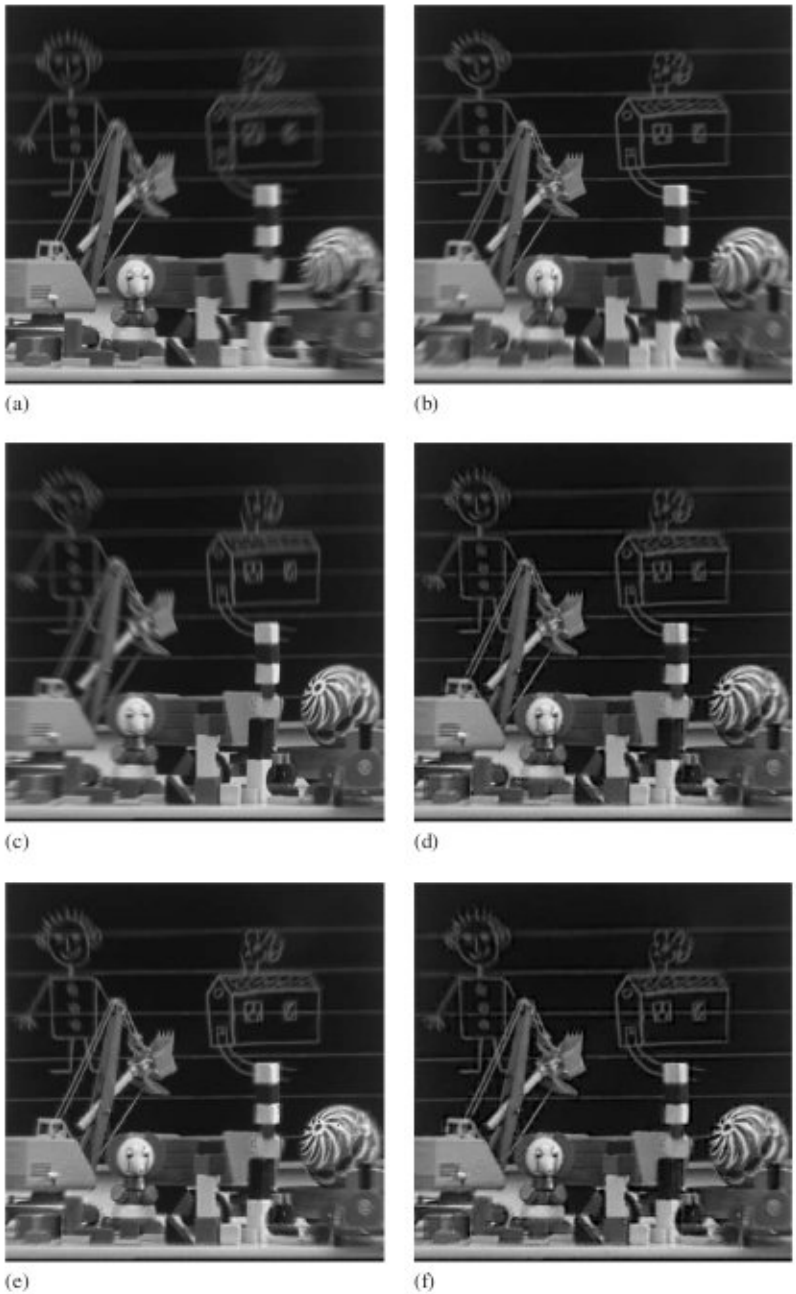


Fig. 5. Original multifocus images and the fused images by different algorithms: (a) Left focused image; (b) center focused image; (c) right focused image; (d) fused image with proposed algorithm; (e) fused image by Haar wavelet; (f) fused image by Heijmans and Goutsias' method.

cannot restore $X(r,c)$ from $G(r,c)$. However, $G(r,c)$ for the image with all parts properly focused may be obtained from various partially focused images as follows. For a set of n multifocus images $X_i, i = 1,\ldots,n$, the gradient images $G_i, i = 1,\ldots,n$ are obtained first. Then, $G_i, i =$

(a) Focus on background (b) Focus on foreground

(c) Fused image with proposed algorithm (d) Fused image by Haar wavelet

(e) Fused image by Heijmans and Goutsias method

Fig. 6. Original multifocus images and the fused images by different algorithms: (a) Focus on background; (b) focus on foreground; (c) fused image with proposed algorithm; (d) fused image by Haar wavelet; (e) fused image by Heijmans and Goutsias' method.

$1, \ldots, n$ are combined into G by taking the maximum gradient value at each position, i.e.

$$G(r,c) = \max\{G_1(r,c), G_2(r,c), \ldots, G_n(r,c)\} \quad \text{for all } (r,c).$$

Thus only the sharply focused regions from the constituent images have their contribution in the *maximum gradient image* G. Let G' denote the *gradient image* obtained from the reconstructed image X'. It is referred to as the *gradient*

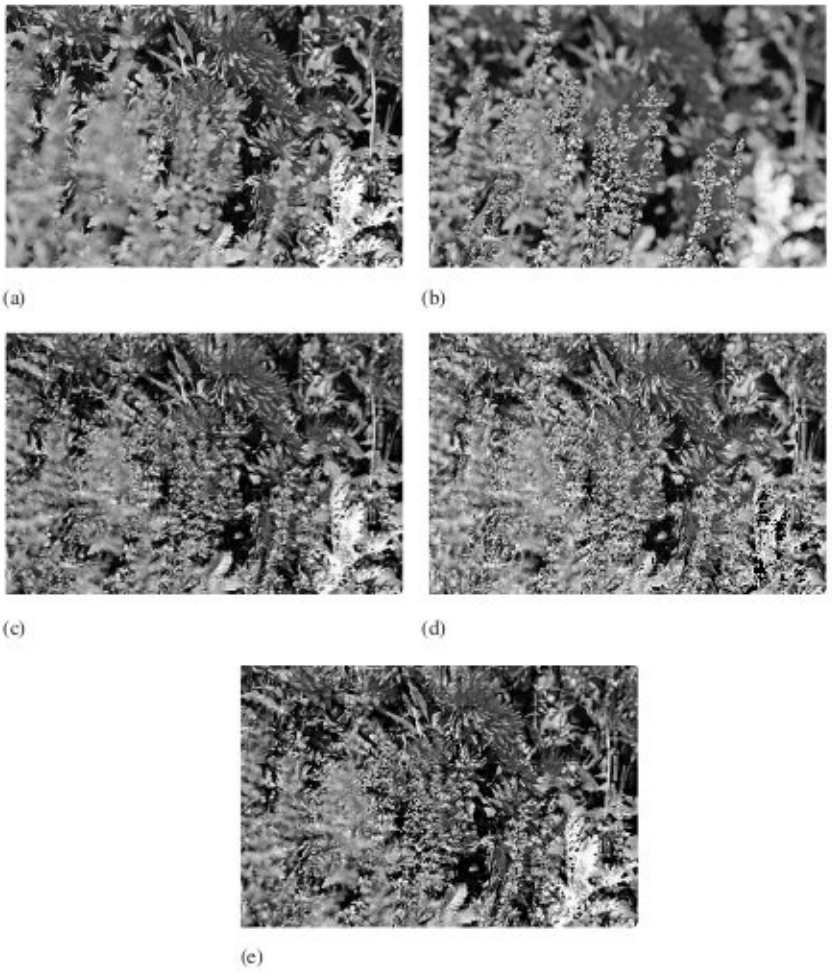


Fig. 7. Original multifocus images and the fused images by different algorithms: (a) Focus on background; (b) focus on foreground; (c) fused image with proposed algorithm; (d) fused image by Haar wavelet; (e) fused image by Heijmans and Goutsias' method.

of fused image. Then, more similar G and G' are, better is the fusion algorithm. The *similarity* S between two images is calculated as

$$S(G,G') = 1 - \frac{\sqrt{\sum (G(r,c) - G'(r,c))^2}}{\sqrt{\sum (G(r,c))^2} + \sqrt{\sum (G'(r,c))^2}}.$$

Hence, for an ideal fused image S approaches the value 1. Similarity between maximum gradient and fused gradient images are listed in Table 1. This table also conforms with the quality measured through manual inspection.

The results obtained by our method are quite good besides the fact that artifact such as blocking effects are noticed in the fused images in some cases (see Fig. 4). But this is a common phenomena in pixel-based image fusion using

Table 1
Similarity between maximum gradient and fused gradient images

Figure	Proposed Algo	Haar wavelet	Heijmans and Goutsius' method
Fig. 4	0.861	0.824	0.853
Fig. 5	0.823	0.835	0.823
Fig. 6	0.929	0.926	0.930
Fig. 7	0.840	0.744	0.835

multiresolution approach and happens due to the fact that error introduced at the topmost level is amplified during reconstruction [6]. In our case, these effects are obvious at the regions where the data is out of focus in all the source images. They are present in the fused images obtained by the other two methods as well.

Apart from the quality of the results, the proposed algorithm has some computational advantages as well. Unlike two other wavelets experimented with, our method ensures that integer pixel values are mapped to integer values only during both analysis and synthesis. This is an useful property for lossless data compression [15]. Secondly, irrespective of the number of times the analysis operators are applied, the range of the values in the scaled images will be same as that of the original multifocus images, say $[0, R]$, and the range of the detail values will be $[-R, R]$. Hence memory-space required during decomposition is fixed. Thirdly, simple arithmetic operations like addition, subtraction and comparison are the only operations used in the method. Other two methods involve division operation and thus they either requires floating point arithmetic or introduces truncation error by using integer arithmetic. Fourthly, due to the non-linear nature of the proposed method, important geometric information (e.g. edges) is well-preserved at lower resolutions. Finally, the method is very fast due to its simplicity. For a set of n, $M \times N$ images, it takes only $O(n \times M \times N)$ computational time. The simplicity of the method and the use of integer arithmetic makes it suitable for chip-level implementation.

Besides this, the non-linear wavelet proposed by us possesses the following invariance properties. Both analysis and synthesis operators are translation invariant in the spatial domain. In the frequency domain, they are grey shift (multiplication) invariant. That means adding (multiplying) a certain value to all pixel values in the original data will result in adding (multiplying) that value to the scaled signal data during analysis [12]. Also, details will not change in case of addition and will get multiplied by that value in case of multiplication. The wavelets possessing these invariance properties, offer better option for image fusion than those which do not possess them [8].

5. Conclusion

In this paper, we have presented a non-linear wavelet constructed by morphological operators and also presented a multifocus image fusion algorithm based on that wavelet. The results are quite impressive considering the fact that the computational cost is very low. The use of simple arithmetic operations makes the method suitable for hardware implementation. However the method suffers from the problem that sometimes artifacts are formed in the fused image. But this is a common problem for all other methods experimented with in this paper. Our method is definitely better than the Haar wavelet method and is at par with Heijmans and Goutsias wavelet method in this respect. These effects are obvious at the regions where the data is out of focus in all the source images. So if the intersection of the defocused regions from all the source images are reduced, then this effect will be minimized.

Acknowledgements

We gratefully acknowledge the sources of data used in our experiments. The images of indoor scene [4] and garden scene [7] are obtained from the websites

http://www.geocities.com/j1hagan/lessons/depthofield.htm and

http://www.photozone.de/4Technique/compose/dof.htm.

The original multifocus images in both sets were unregistered and the registration process was done by Mr. Buddhajyoti Chattopadhyay. The registered multifocus images in Figs. 5 and 6 were obtained from the site

http://www.ece.lehigh.edu/SPCRL/IF/multifocus.htm.

References

[1] W. Seales, S. Dutta, Everywhere-in-focus image fusion using controllable cameras, Proceedings of SPIE 2905, 1996, pp. 227–234.

[2] I. Bloch, Information combination operators for data fusion: a review with classification, IEEE Trans. SMC: Part A 26 (1996) 52–67.
[3] H.A. Eltoukhy, S. Kavusi, A computationally efficient algorithm for multi-focus image reconstruction, Proceedings of SPIE Electronic Imaging, June 2003.
[4] S. Mukhopadhyay, B. Chanda, Fusion of 2d gray scale images using multiscale morphology, Pattern Recognition 34 (2001) 1939–1949.
[5] P.J. Burt, R.J. Lolezynski, Enhanced image capture through fusion, in: Proceedings of the Fourth International Conference on Computer Vision, Berlin, Germany, 1993, pp. 173–182.
[6] H. Li, B. Manjunath, S. Mitra, Multisensor image fusion using the wavelet transform, Graph. Models Image Process. 57 (3) (1995) 235–245.
[7] X. Yang, W. Yang, J. Pei, Different focus points images fusion based on wavelet decomposition, Proceedings of Third International Conference on Information Fusion, vol. 1, 2000, pp. 3–8.
[8] Z. Zhang, R.S. Blum, Image fusion for a digital camera application, Conference Record of the Thirty-Second Asilomar Conference on Signals, Systems and Computers, vol. 1, 1998, pp. 603–607.
[9] S. Mallat, A Wavelet Tour of Signal Processing, Academic Press, London, 1998.
[10] W. Sweldens, The lifting scheme: a new philosophy in biorthogonal wavelet constructions, in: A. Laine, M. Unser, (Eds.), Wavelet Applications in Signal and Image Processing, vol. III, Proceedings of SPIE, vol. 2569, 1995, pp. 68–79.
[11] J. Goutsias, H.J. Heijmans, Nonlinear multiresolution signal decomposition schemes, Part 1: morphological pyramids, IEEE Trans. Image Process. 9 (November 2000) 1862–1876.
[12] H.J. Heijmans, J. Goutsias, Multiresolution signal decomposition schemes, Part 2: morphological wavelets, IEEE Trans. Image Process. 9 (November 2000) 1897–1913.
[13] B. Chanda, D.D. Majumdar, Digital Image Processing and Analysis, Prentice-Hall of India, New Delhi-110001, 2000.
[14] L.G. Roberts, Machine perception of three dimensional solids, in: J.T. Tippet (Ed.), Optical and Electro-optical Information Processing, MIT Press, Cambridge, MA, 1965.
[15] A.R. Calderbank, I. Daubechies, W. Sweldens, B.L. Yeo, Wavelet transforms that map integers to integers, Appl. Comput. Harmonic Anal. 5 (1998) 332–369.