

## Development of Video Fusion Algorithm at Frame Level for Removal of Impulse Noise

<sup>1</sup>P.Nalini, <sup>2</sup>P.Lakshmi <sup>3</sup>P.Sri Hari, <sup>4</sup>Ch.Gayatri,  
<sup>1</sup>M.Tech,SSP DIET College of Engineering,Visakhapatnam.  
<sup>2,3,4</sup>Dept. of ECE,DIET College of Engineering, Visakhapatnam.

---

**Abstract** — Satellite communications is growing rapidly in various fields like GPS. In general satellite gives the information in the form of images and videos. This data is often corrupted during acquisition, transmission or due to faulty memory locations in hardware. This creates loss of information. The noise density varies depending on various factors namely reflective surfaces, atmospheric variations, noisy communication channels etc. The main noise effecting the satellite data is Impulse noise. In this paper an algorithm is developed to remove the impulse noise in the video captured by satellites. In this paper a video captured by different sensors at same time instant with impulse noise are considered and as order statistics filters exhibit better performance, these noisy videos are filtered individually using non linear filtering algorithms namely, Vector Median Filter (VMF), Rank Conditioned VMF (RCVMF), Rank conditioned and Threshold VMF (RCTVMF) and Absolute Deviation VMF (ADVMF) . These filtered videos from individual filtering methods are combined to a single video called video fusion, which retains the important features of the frames in the video and thereby not only removing the noise but also preserve the finer details of image hence increase the quality of the image . In this paper the filtered frames from the captured video are fused into a single video by using the video fusion technique which is based on the quality assessment in spatial domain. For future enhancement in the video this fused video is again filtered using absolute deviation VMF which gives best result than the fused and filtered videos. The performance evaluation of the filtered and the fused video with respect to the original video is done using mean square error(MSE), peak signal to noise ratio(PSNR) and structural similarity index(SSIM).

**Keywords** — Median Filter, RCVMF, RCTVMF, ADVMF, Video Fusion, Order Static filters, SSIM

---

### I. INTRODUCTION

Noise is undesired information that contaminates an image. Noise appears in an image from various sources. The digital image acquisition process converts an optical image into a continuous electrical signal. This electrical signal sampled, is primary process by which noise appears in digital image. Video is the continuous form of image data called frames. There are several ways through which noise can be introduced into data, depending on how the data is retrieved. This is the main problem in remote sensing applications. Satellite videos, containing the noise signals and lead to a distortion and not being able to understand. So to study it properly, requires the use of appropriate filters to limit or reduce much of the noise. It helps the possibility of better interpretation of the content of the data. The videos captured by multiple sensors can be differently noised depending on the proximity to the object, environmental disturbances and sensor features and they may be noised to same intensity some times. There are different types of noises are added to the data depending on the nature of sensor, application, hardware etc. Gaussian noise, uniform noise, impulse noise, gamma noise, exponential noise etc. are some of the noise models in image processing. Impulse noise is one of the common problems in digital data. In this paper an attempt has been made to remove impulse noise from videos captured by different sensors. The impulse noise can be removed by using nonlinear filters. The median filter is mostly used nonlinear filter to remove impulse noise. However the direct extension of the median filter to color video is not a straight forward one. The ordering of the vector pixel data is carried by considering suitable distance measure. The vector pixels in a kernel are ordered depending on the sum of the distances between each vector pixel and the other vector pixels in a window or kernel. The vector pixel with smallest sum of distances forms the vector median. In most of the imaging applications, for example remote sensing, the video captured is converted in to images by different sensors and are combined to a single image, which retains the important features of the images from individual sensors , hence the video itself. Only removal of noise is not sufficient for further processing in case remote sensing applications. The filtered images must be fused into a single image which is most suitable than individual images from the filters. This task accomplished by a process called image fusion. video fusion is defined as the process of combining substantial information from several sensors using mathematical techniques in order to create a single composite signal that will be more comprehensive and thus,

more useful for a human operator or other computer vision tasks. Current technology in image sensors offers a wide variety of information that can be extracted from an observed scene.

## II. ORDER STATIC FILTER FOR COLOR VIDEO PROCESSING

### A. Vector Median Filter

In the vector median filter for the ordering of vectors in a particular kernel or mask a suitable distance measure is chosen. The vector pixels in the window are ordered on the basis of the sum of distances between each vector pixel and the other pixels in the window. The sum of the distances is arranged in ascending order and then the same ordering is associated with vector pixels. The vector pixel with the smallest distances is the vector median pixel. The vector median filter is represented as

$X_{VMF} = \text{Vector Median}(a_A)$

Where for the image IM with window size  $W$ ,  $L_k$  represents the sum of distances from  $K^{\text{th}}$  pixel to all other pixels in the kernel and where  $(1 \leq i \leq W^2)$  and  $a_A(k)$ ,  $a_A(p)$  the vectors,  $L_k$  is the distance measure is given by the city block distance which is more suited to non correlated noise. The ordering may be illustrated as

$$L_1 \leq L_2 \leq L_3 \leq \dots \leq L_w^2$$

$$\sum_{i=w-\frac{(w-1)}{2}}^{m-\frac{(w-1)}{2}} \sum_{j=w-\frac{(w-1)}{2}}^{n-\frac{(w-1)}{2}} \left( \left( \left( \sum_{x=-\frac{(w-1)}{2}}^{\frac{w-1}{2}} \sum_{y=-\frac{(w-1)}{2}}^{\frac{w-1}{2}} a_A = \text{IM}(x+i, y+j) \right) \sum_{k=1}^{w^2} \sum_{p=1}^{w^2} L_k = (a_A(k) - a_A(p)) + L_Q \right) \right)$$

Where  $A = 1$  when  $i = j = w - \frac{(w-1)}{2}$

$A = 1+A$  when  $x + y = -2$

Where  $L_Q = 0$  when  $k = m = 1$

$L_Q = L_K$  else

And this implies the same ordering to the corresponding vector pixels i.e.

$$a_A(1) \leq a_A(2) \leq a_A(3) \dots \leq a_A(w^2)$$

Since the vector pixel with the smallest sum of distances is the vector median pixel, it will correspond to rank 1 of the ordered pixels i.e.,

$$X_{VMF} = a_A(1)$$

### B. Rank Conditioned Vector Median Filter

The rank conditioned vector median filter improves the performance of the vector median filter. The vector median of the kernel replaces the central pixel when the rank of the central pixel is greater than a predefined rank of a healthy vector pixel inside the window. To find out a rank of the healthy vector pixel the code simulating RCVMF is executed on a noiseless image. Then, the mean value of the obtained ranks of the central vector pixel is calculated. This value is rounded off to a whole number, and it is considered to be rank of the healthy vector pixel of a kernel. Then this rank is used for simulations at various noise percentages. The rank of a healthy vector pixel is given as

$$M_{\theta} = \frac{\left( \sum_{i=w-\frac{(w-1)}{2}}^{m-\frac{(w-1)}{2}} \sum_{j=w-\frac{(w-1)}{2}}^{n-\frac{(w-1)}{2}} \left( R_c = \left( \text{rank} \left( \sum_{x=-\frac{(w-1)}{2}}^{\frac{w-1}{2}} \sum_{y=-\frac{(w-1)}{2}}^{\frac{w-1}{2}} a_A = \text{im}(x+i, y+j) \right) + R_c \right) \right) \right)}{(m \times n) - \frac{(w-1)}{2}}$$

Where  $R_c = 0$  when  $i = j = w - \frac{(w-1)}{2}$

The vector median value  $X_{VMF}$  can be calculated as

$$\sum_{i=w-\frac{(w-1)}{2}}^{m-\frac{(w-1)}{2}} \sum_{j=w-\frac{(w-1)}{2}}^{n-\frac{(w-1)}{2}} \left( \left( \left( \sum_{x=-\frac{(w-1)}{2}}^{\frac{w-1}{2}} \sum_{y=-\frac{(w-1)}{2}}^{\frac{w-1}{2}} a_A = \text{IM}(x+i, y+j) \right) \sum_{k=1}^{w^2} \sum_{p=1}^{w^2} L_k = (a_A(k) - a_A(p)) + L_Q \right) \right)$$

$L_Q$

Where  $A = 1$  when  $i = j = w - \frac{(w-1)}{2}$

$A = 1+A$  when  $x + y = -2$

Where  $L_Q = 0$  when  $k = m = 1$

$L_Q = L_K$  else

The rank conditioned vector median filter can be expressed as:

$$\begin{aligned} X_{RCVMF} &= X_{VMF}, & \text{if } R_c > M_\theta \\ &= X_c & \text{otherwise} \end{aligned}$$

Where  $R_c$  is a rank of the central vector pixel and  $c=5$  for the window length of  $w = 3$  and so on and  $M_\theta$  is the predefined healthy vector pixel rank inside the window, where  $X_c$  is the center vector pixel of the kernel. Image details are preserved better if the rank chosen is large. However this is at the cost of passing more impulses are removed to the maximum extent with a smaller value of the rank but blurring effect is more pronounced in the restored image.

**C. Rank Conditioned Threshold Vector Median Filter**

The rank-conditioned and threshold vector median filter aims to further enhance the RCVMF by incorporating an additional test – a distance threshold for the detection of impulses. In RCVMF a central vector having greater than the predefined rank implies a corrupt vector. However, it may not be true always, because the vectors may be close as per the distance measure. Hence another criterion  $\Theta$  is taken into account. It is the distance between the central vector pixel and the vector pixel corresponding to the predefined rank. To find out the value of this pre-determined distance threshold  $\Theta$ , the code simulating RCTVMF is executed on a noiseless image. Then the mean of the obtained  $\Theta$  values is calculated and used for the simulations at various noise percentages. This value is used in the program to help select only the corrupt central vector pixels.

The distance threshold  $\Theta$  is calculated as follows

$$L_Q = \sum_{i=w-\frac{w-1}{2}}^{m-\frac{w-1}{2}} \sum_{j=w-\frac{w-1}{2}}^{n-\frac{w-1}{2}} \left( \left( \sum_{x=-\frac{w-1}{2}}^{\frac{w-1}{2}} \sum_{y=-\frac{w-1}{2}}^{\frac{w-1}{2}} a_A = IM(x+i, y+j) \right) \sum_{s=1}^{w^2} L_p = \left( (a_A((w^2+1)/2) - a_A(s)) \right) \right)$$

Where	A = 1	when $i = j = w - \frac{w-1}{2}$
	A = 1+A	$x + y = -2$
Where	$L_Q = 0$	when S = 1
	$L_Q = L_p$	else
Where	p = 1	when $i = j = w - \frac{w-1}{2}$
	p = p + 1	else

$$\text{Where } \Theta = L_p / \left( (m \times n) - \frac{w-1}{2} \right)$$

The rank threshold value is calculated as

$$M_\theta = \frac{\sum_{i=w-\frac{w-1}{2}}^{m-\frac{w-1}{2}} \sum_{j=w-\frac{w-1}{2}}^{n-\frac{w-1}{2}} \left( R_c = \left( \text{rank} \left( \sum_{x=-\frac{w-1}{2}}^{\frac{w-1}{2}} \sum_{y=-\frac{w-1}{2}}^{\frac{w-1}{2}} a_A = \text{im}(x+i, y+j) \right) + R_c \right) \right)}{(m \times n) - \frac{w-1}{2}}$$

Where $R_c = 0$	when $i = j = w - \frac{w-1}{2}$
-----------------	----------------------------------

On the basis of the above information, the filter has the following form:

$$\begin{aligned} X_{RCTVMF} &= X_{VMF}, & \text{if } R_c > M_\theta \text{ and } L_p > \Theta \\ &= X_c & \text{otherwise} \end{aligned}$$

The distance  $L_p$  is usually very big if impulses are present at the center of the window. If the value of threshold is reduced, more and more number of the central vector pixels will be replaced by the vector median and the filter will resemble the RCVMF.

**D. Absolute Deviation Vector Median Filter**

In this filter, the impulse noise detection mechanism does not require the distance calculation and subsequent ordering of the vectors of a kernel. The algorithm deals with the difference values of the red (R) and the green (G) intensities denoted by  $\Omega_{RGi}$  (say), and the difference values of the green (G) and blue (B) intensities denoted by  $\Omega_{GBi}$  (say), (where  $1 \leq i \leq N, N=9$ ).

In a  $3 \times 3$  kernel, it has been observed empirically that  $\Omega_{RGi}$  and  $\Omega_{GBi}$  values closely correspond to each other. Thus the mean absolute deviation  $\mathbf{D}'_{RG}$  and  $\mathbf{D}'_{GB}$  i.e. the mean of  $\mathbf{D}_{RGj}$  and  $\mathbf{D}_{GBj}$  (where  $1 \leq j \leq N$ , and  $j \neq c$ ,  $c = (N+1)/2$ ,  $N = 9$ ) has small values.  $\mathbf{D}_{RGj}$  and  $\mathbf{D}_{GBj}$  are the absolute deviation values of  $\Omega_{RGj}$  and  $\Omega_{GBj}$  from  $\Omega'_{RG}$  and  $\Omega'_{GB}$  respectively.  $\Omega'_{RG}$  and  $\Omega'_{GB}$  denote the mean of  $\Omega_{RGj}$  and  $\Omega_{GBj}$  (where  $1 \leq j \leq N$ , and  $j \neq c$ ,  $c = (N+1)/2$ ,  $N = 9$ ). The absolute deviation of the central vector  $\Omega_{RGc}$  and  $\Omega_{GBc}$  values from  $\Omega'_{RG}$  and  $\Omega'_{GB}$  is obtained respectively as  $\mathbf{D}_{RGc}$  and  $\mathbf{D}_{GBc}$ .

If the absolute deviation  $\mathbf{D}_{RGc}$  and  $\mathbf{D}_{GBc}$  of the central vector pixel exceeds the value of  $\mathbf{D}'_{RG}$  or  $\mathbf{D}'_{GB}$  respectively for a  $3 \times 3$  kernel, the central vector pixel is to be replaced by the vector median of the kernel. The algorithm is represented as follows.

In a  $3 \times 3$  kernel,

*Step 1* Find the difference values of red(R) and the green (G) intensities denoted by  $\Delta_{RGi}$  and the difference values of the green (G) and blue (B) intensities denoted by  $\Delta_{GBi}$

$$\begin{aligned} \Delta_{RGi} &= X(i, R) - X(i, G) \\ \Delta_{GBi} &= X(i, G) - X(i, B) \quad \text{Where } i=1 \dots 9 \end{aligned}$$

*Step 2* Calculate the mean of  $\Delta_{RGi}$  and  $\Delta_{GBi}$  denoted by  $\Delta^1_{RGi}$  and  $\Delta^1_{GBi}$ . Where  $i=1,2 \dots 9$  and  $i \neq 5$ .

*Step 3* Calculate the absolute deviation between  $\Delta$  and  $\Delta^1$ .

$$\begin{aligned} \mathbf{D}_{RGi} &= |\Delta_{RGi} - \Delta^1_{RGi}| \\ \mathbf{D}_{GBi} &= |\Delta_{GBi} - \Delta^1_{GBi}| \quad \text{where } i=1,2, \dots 9 \end{aligned}$$

*Step 4* Calculate the mean of  $\mathbf{D}_{RGi}$  and  $\mathbf{D}_{GBi}$  denoted as  $\mathbf{D}^1_{RGi}$  and  $\mathbf{D}^1_{GBi}$ . Where  $i=1, 2 \dots 9$  and  $i \neq 5$ .

*Step 5* Now  $\mathbf{D}_{RGc} = |\Delta_{RGc} - \Delta^1_{RGi}|$   
 $\mathbf{D}_{GBc} = |\Delta_{GBc} - \Delta^1_{GBi}|$  where  $i=1,2, \dots 9$ ,  $i \neq 5$  and  $c$  is the central vector.

*Step 6* If  $\mathbf{D}_{RGc} > \mathbf{D}^1_{RGi}$  OR  $\mathbf{D}_{GBc} > \mathbf{D}^1_{GBi}$  where  $i=1,2 \dots 9$  and  $i \neq 5$ . Central vector is corrupted; hence central vector is replaced by Vector Median of kernel.

### III. QUALITY ASSESSMENT OF SPATIAL DOMAIN

Quality assessment of images has been successfully employed in the authentication area, such as iris and fingerprint verification, which is used to evaluate the quality of the captured images. Fingerprint quality is usually defined as a measure of the clarity of the ridge and valley structures. In the multi-focus image fusion, it is the first step to investigate whether the region of the image is clarity or not, so we incorporate the quality assessment into multi-focus image fusion algorithm. Generally speaking, the quality assessment can be classified into two ways, namely the frequency domain assessment and the spatial domain assessment. We incorporate the latter one into this algorithm.

In order to assess the image quality in a local region, we partition a given image into a lattice of blocks of size  $b \times b$ . For each block B, let  $\mathbf{g}_s = (g_s^x, g_s^y)$  denote the gradient of the gray level intensity at site  $s \in B$ . The covariance matrix of the gradient vectors for all  $b^2$  sites in this block is given by

$$J = \frac{1}{b^2} \sum_{s \in B} \mathbf{g}_s^T \mathbf{g}_s = \begin{bmatrix} j_{11} & j_{12} \\ j_{21} & j_{22} \end{bmatrix}$$

The above symmetric matrix is positive semi definite with Eigen values

$$\begin{aligned} \lambda_1 &= \frac{1}{2} \left( \text{trace}(J) + \sqrt{\text{trace}(J)^2 - 4 \det(J)} \right) \\ \lambda_2 &= \frac{1}{2} \left( \text{trace}(J) - \sqrt{\text{trace}(J)^2 - 4 \det(J)} \right) \end{aligned}$$

Where  $\text{trace}(J) = j_{11} + j_{22}$ ,  $\det(J) = j_{11}j_{22} - j_{12}^2$  and  $\lambda_1 \geq \lambda_2$ . The quality assessment defined as  $\lambda = \lambda_1 - \lambda_2$ . This measure reflects the clarity of the local region. The clearer the local region is, the bigger the measure is.

### IV. RESULTS

In general the data captured by different sensors are corrupted with noise. The noise from the data must be removed for further processing. Once noise removed from the images taken from video, these images are combined into a single image or frame and then forms in to efficient video. Because, individual sensor cannot provide all the information. Especially in case of remote sensing applications it is necessary to combine the information from all sensors. This is done by video fusion technique. The video fusion technique enables the combination of all individual frames of the captured video into a single frame without losing any original information.

In this paper data captured by different sensors are considered. It is assumed that the videos captured are contaminated with impulse noise with different noise density (10%-60%). The videos are filtered

individually using different filtering algorithms namely, Vector Median Filter (VMF), Rank Conditioned Vector Median Filter (RCVMF), Rank conditioned and Threshold Vector Median Filter (RCTVMF) and Absolute Deviation Vector Median Filter (ADVMF). Each filter removes the impulse noise but of all these filtered videos ADVMF filtered video is best .So fusion is done using ADVMF filtered videos are fused for further processing The performance of each filter can be evaluated by using image quality measures can also be used to evaluate the video. The common image quality measures are mean square error, peak signal to noise ratio and structural similarity index. The fused video is compared with original video at frame level using these image quality measures. The performance of each filter and fusion of these videos can be evaluated by image quality measures.

The figure 1 shows the Mean Square Error of each filter and fused video at different noise densities. The figure 2 shows the peak signal to noise ratio of each filter and fused videos at different noise densities , the figure 3 shows the structural similarity index of each filter and fused video at different noise densities and figure 4 shows the Block diagram. Each filter specified with different colors in figures. Here TABLE I depicts the data MSE, PSNR, SSIM for a filtered video which is contaminated with 60% impulse noise are represented.

TABLE I  
Comparison of different filtering Methods and Fusion Technique for 0.6% Impulse Noise

Noise Density of 60%			
Method	MSE	PSNR	SSIM
VMF	3807.41	12.3243	0.8601
RCVMF	3900.62	11.9614	0.8541
RCTVMF	3803.83	12.3285	0.8614
ADVMF	815.555	19.0167	0.9287
FUSION	326.228	22.9966	0.9409

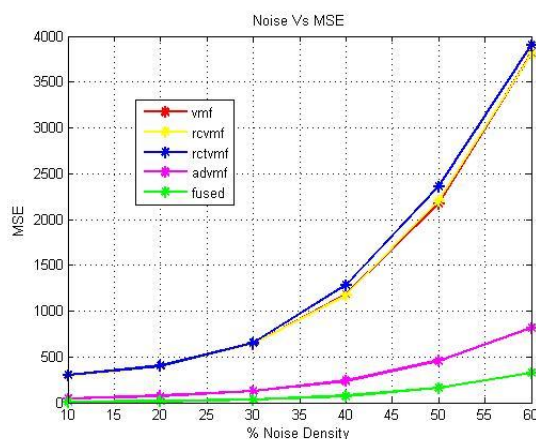


Fig. 1 Comparison of MSE with Different Noise intensities for the filtering techniques

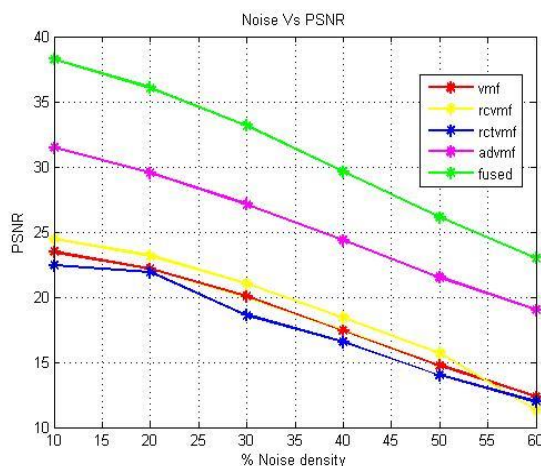


Fig. 2 Comparison of PSNR with Different Noise intensities for the filtering techniques

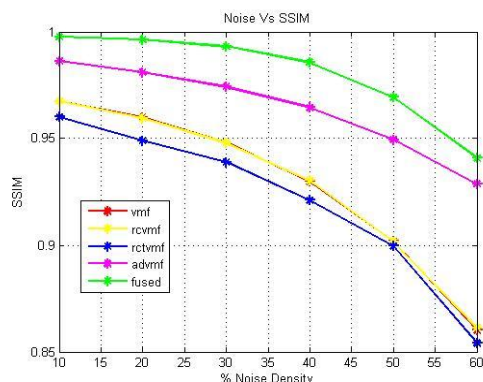


Fig. 3 Comparison of SSIM with Different Noise intensities for the filtering techniques

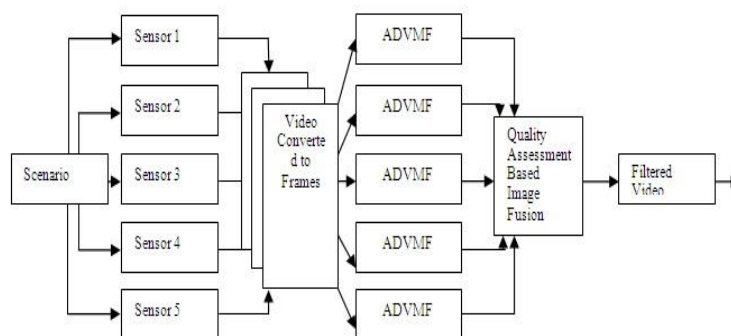


Fig 4 Block Diagram of Multi Sensor Video Fusion at Frame Level Using QA Technique

## V. CONCLUSION

Digital video is often corrupted by impulse noise captured by sensors. The impulse noise in a video captured by sensor depends on the combination of an illumination source and the reflection or absorption of energy from that source. This noise should be removed from the data for further processing. There are different algorithms for the removal of impulse noise. For impulse noise, median filter is the best choice. The different filtering techniques for removal of impulse noise are presented. In this paper vector median filter, rank conditioned vector median filter, rank conditioned & threshold vector median filter and absolute deviation vector median filter are analyzed by applying to remote sensing video for removal of impulse noise. But in some applications only removal of noise from data is not sufficient. Instead of using single sensor and single filtering technique, we are using different sensors for capturing the video and different filtering techniques are applied simultaneously for removal of noise. By considering quality assessment of each filtered video captured by different sensors, the videos are fused which results in more noise free video than any individual filtered video. The filtered videos are fused at frame level based on the quality assessment in spatial domain for removal of impulse noise. Now this fused video at frame level is again filtered using absolute deviation VMF which performs better than fused image. The performance evaluation of each filter and fusing technique is done by using image quality measures like mean square error, peak signal to noise ratio and structural similarity index.

## REFERENCES

- [1] Nikolaos Mitianoudis, Tania Stathaki, Pixelbased and Region-based Image Fusion schemes using ICA bases, Communications and Signal Processing group, Preprint submitted to Elsevier Science, 13th Dec 2007.
- [2] HyvÄarinen, P. O. Hoyer, and E. Oja. Image denoising by sparse code shrinkage. In S. Haykin and B. Kosko, editors, *Intelligent Signal Processing*. IEEE Press, 2001.
- [3] P. Hill, N. Canagarajah, and D. Bull. Image fusion using complex wavelets. In *Proc. 13<sup>th</sup> British Machine Vision Conference*, Cardiff, UK, 2002.
- [4] M. Vetterli and C. Herley, "Wavelets and filter banks: theory and design," *IEEE Trans Signal Processing*, vol. 40, pp. 2207-2232, September 1992.
- [5] P. Burt and R. Lölczynski, "Enhanced image capture through fusion," in *Proc. the Fourth International Conference on Computer Vision*, pp. 173- 182, 1993.