2012 IEEE Ninth International Conference on Advanced Video and Signal-Based Surveillance

Combining Infrared and Visible Images using Novel Transform and Statistical Information

Jeongmin Bae
Dept. Visual Information Processing
Korea University
Seoul, Korea
jmbae@ispl.korea.ac.kr

Bonwha Ku
School of Electrical Engineering
Korea University
Seoul Korea
bhku@ispl.korea.ac.kr

David K. Han
Ocean Engineering and Marine Systems Team
Office of Naval Research
Arlington, VA, USA
ctmkhan@gmail.com

Hanseok Ko
School of Electrical Engineering
Korea University
Seoul Korea
hsko@korea.ac.kr

***Abstract*—This paper proposes a novel combining method of infrared (IR) and visible images based on a Discrete Wavelet Frame (DWF) approach. In contrast to existing methods, IR image is transformed first using statistical information of the visible image to emphasize relevant information. In a multi-scale domain, we then assign appropriate weights to each pixel of sub-band approximation images through pixel level weighted average for emphasizing relevant information of the IR image while keeping texture information of the visible image. Representative experiments show that the proposed method outperforms exiting methods in image quality.**

Keywords- Image fusion, Image processing, LWIR

I. Introduction

In image-based surveillance, both video and Infrared (IR) cameras are widely used. In particular, Long Wave IR cameras can capture images independent of illumination and even in foggy conditions to a certain extent. Since its image is composed of thermal distributions, live animals or persons hidden in vegetation stand out from the background prominently. However, LWIR has limitations of capturing color or image details of fine textures. A visible camera, on the other hand, has complimentary capabilities of effectively capturing color and fine texture information in a scene. It, however, requires sufficient level of illumination to be able to capture useful images. Due to their complimentary nature of capabilities and limitations, there exist a particular focus of interest in the fusion of visible and Infrared (IR) sensors.

Currently, there exist many algorithms to generate a single image from visible and IR sensors. Since Mallat suggests [1] applying Discrete Wavelet Transform (DWT) for fusion algorithm, research trend followed to use DWT or its improved analysis method. Most of these approaches were based on combining source images which is decomposed by a multi scale decomposition transform. A combined image is reconstructed by an inverse transform as a result. Rockinger and Fechner proposed [2] a fusion algorithm suitable for a series of sequential images, such as a movie, using Discrete Wavelet Frame (DWF) to overcome the shift variance property of DWT. Miaw and Wang proposed an algorithm [4] that uses a contourlet transform which is efficient to represent the singularity of linear and curve in analysis for preserving edges and texture information. Hill et al. proposed a method [5] using a Dual Tree Complex Wavelet Transform (DT-CWT) to take advantage of computational efficiencies and edge preserving property. These algorithms focus more on the analysis method than combination rules to prevent loss of edge information. Zhang and Blum mentioned window-based activity measure using weighted average (WA-WBA) for a combination rule [6]. They use simple combination rules like average or weighted average for coarse image level combining. This brings on a contrast reduction problem of the fused image because typical surveillance images taken by an IR sensor exhibits small variance by temperature. This is a significant problem because human eyes are more sensitive to contrast changes than shape changes such as edges [?]. For night-time surveillance images, fusion of an IR image to a video image doesn't reduce the contrast of the video image significantly. However, for day-time surveillance, an added process is needed to mitigate the contrast reduction problem.

The key objective of visible and IR image fusion is to capture and represent the main information contents into a single image. To achieve this objective, we transformed an IR image to emphasize major information of IR and suppress rest of the data. We calculate weights for combining the transformed IR image and the visible image for the fusion process. The main contribution presented in this paper is an effective IR transform method and a novel weighted average method of approximation level combining rules for a multiscale decomposition based image fusion.

978-0-7695-4797-8/12 $26.00 © 2012 IEEE
DOI 10.1109/AVSS.2012.24

(a) LWIR (b) visible (c) fusion result

Figure 1. Contrast reducing problem of conventional weighted average rule for image fusion

Outline of this paper is as follow: In Section 2 we present the proposed IR image transform method and the novel weighted sum rule. In Section 3, we compare the proposed method to a conventional algorithm with a discussion. Conclusion is included in Section 4.

II. PROPOSED FUSION ALGORITHM

Objects of interest in surveillance applications, such as human, animals, or a vehicle exhibits higher IR intensities than neighboring regions, because they are typically at higher temperature than the surroundings. Therefore, these can be easily detected by a LWIR camera which senses heat sources independent of illumination conditions. Even if those objects are concealed in vegetation, an IR camera may detect them from the background.

Our algorithm is based on a multi scale decomposition fusion method. A block diagram of the proposed algorithm is shown at Fig. 2.

A. IR image transformation

IR image transformation is a process that divides the IR image into two regions and emphasizes which has higher pixel intensity regions. To divide IR images, we use a mean value of visible image as a threshold. Transformation equation is (1). x_{IR} and x_{IRt} mean pixel elements of transformed IR image and input IR image. μ_{CCD} is the mean value of the visible image. We add 1 to prevent dividing by zero for weight calculation.

$$x_{IRt} = \begin{cases} \{(x_{IR} - \mu_{CCD}) / \mu_{CCD} + 1\}^2 & \text{where } x_{IR} \le m_{CCD} \\ \{(x_{IR} - \mu_{CCD})\mu_{CCD} + 1\}^2 & \text{where } x_{IR} > m_{CCD} \end{cases} \qquad (1)$$

The IR image is transformed to emphasize the regions of higher intensities while the regions of low IR intensities are de-emphasized accordingly. As shown in Fig. 3, a man on the IR image consisting of high intensity values by the body temperature is emphasized in the combined image while the trees, leafs, road with low intensity values are de-emphasized.

B. Image decomposition

The image fusion algorithms in surveillance applications have to deal with sequential images. Due to shift variance property, DWT is not suitable for a sequenced image fusion. In [2], Rockinger proposed an approach based on DWF which has a property that is shift invariant. Rockinger showed that a fusion method which uses the DWF method is better than the other methods with respect to stability and consistency. We have selected DWF for the multi resolution decomposition method.

DWF decomposes an image into 4 frequency bands: low-low (LL), low-high (LH), high-low (HL) and high-high (HH). LH, HL and HH band provide detail image information such as edges and textures while LL band only retains low frequency components of the image. High bands are denoted as “detail image” in this paper. And LL band is denoted as “approximation image.”

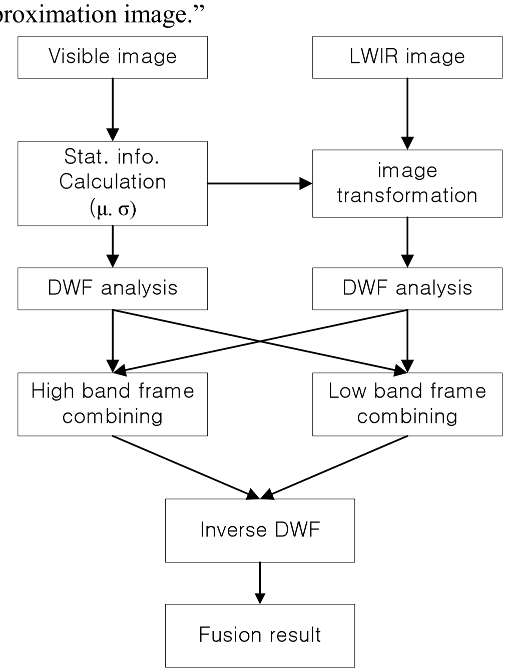


Figure 2. Block diagram of proposed algorithm

(a) LWIR (b) visible (c) Transformed LWIR

Figure 3. LWIR image transformation

C. Detail image combining

Because of different physical meaning, detail and approximation images are combined with different methods [10].

Magnitudes of detail image DWF coefficients correspond to sharp intensity changes. For detail image combination we employ the "choose-max" scheme. which picks the larger coefficients of the IR or the CCD images.

D. Coarse image combining

Basically, our approach was taking ratio each input pixel values and sum of each pixel value to get weight. But while we take experiments, we found the problem that bright light s like a car's headlights occur lens flare effects in visible images as Fig. 4(a). Visible image regions which have lens flare effect cover same regions of IR image. We have changed weight rule as like (3) to solve this problem by applying standard deviation as like (4). The reason why we use standard deviation is that visible with flare effect has been occurred by strong light which reduces aperture. So standard deviation of visible is much larger than IR's one because IR sensor is not effected by light. x_{ccd} and x_{irt} is pixel element of visible and transformed IR images. w_{ccd} and w_{ir} mean pixel weight on visible image and IR image.

$$\begin{aligned} w_{CCD} &= x_{CCD} / (x_{CCD} + x_{IRt} \cdot \alpha) \\ w_{IR} &= x_{IRt} \cdot \alpha / (x_{CCD} + x_{IRt} \cdot \alpha) \end{aligned} \qquad (3)$$

$$\alpha = \begin{cases} C & \begin{cases} if \quad x_{CCD} > C_{IR} \cdot \sigma_{IR} \\ if \quad x_{IR} > C_{CCD} \cdot \sigma_{CCD} \end{cases} \\ 1 & \text{otherwise} \end{cases} \qquad (4)$$

We used standard deviation to reduce influence of lens flare effect and to reflect IR information as like (4). C is user-defined value to control weight on IR for increasing IR pixel value. It causes reducing proportion of x_{ccd} on final weight. C_{IR} and C_{CCD} are user defined constants.

After taking the weights for each visible and transformed IR image, approximation image fusion result, $x_{f,}$ is the sum of each weighted pixel value from the visible and the transformed IR images as in (5).

$$x_f = w_{CCD} x_{CCD} + w_{IR} x_{IRt} \qquad (5)$$

The resultant overall image is of combined detail image and approximation image by inverse DWF.

III. EXPERIMENT

In order to evaluate the performance of the proposed method, we have developed a dataset by taking various scenes; day, night, strong light, snowy. The snowy scene is taken at daytime. Cameras that we use for the experiment is SCB-9051 as the IR sensor and SCO-2120R as the visible sensor from Samsung. The dataset consists of 60 images. We have selected 15 images from each scene. All IR images are registered to visible images. Border regions are padded with corresponding visible image regions.

Generally, two types of quality measures can be used to

(a) visible image

(b) LWIR image

Figure. 4 Lens flare problem on visible image evaluate performance of an image fusion. One type considers features of the fusion result and the input source. The other type is extracting features from the fusion result only. We evaluate experiment results with the both types. The first measure is global contrast factor (GCF). It is a measure of global contrast of a single image. Matkovic et al. proposed this indicator [3] to quantify human perception on contrast. We have selected this measure because the objective of the proposed algorithm is to develop a natural fusion suitable for human perception. As shown in (6), GCF is measured by weighted sum of average local contrast C_i which consists of the sum of the absolute values of each 4-way pixel difference lc_i in a local region. w and h are width and height of the local region. $L(x,y)$ means the pixel value of the location (x,y) on an image.

$$
\begin{aligned}
&GCF = \sum_{i=1}^{N} w_i * C_i \quad, Ci = \frac{1}{w*h}\sum_{i=1}^{w*h}\frac{1}{lc_i} \\
&lc_i = |L(x,y) - L(x-1,y)| + |L(x,y) - L(x+1,y)| \\
&+ |L(x,y) - L(x,y-1)| + |L(x,y) - L(x,y+1)|
\end{aligned}
\tag{6}
$$

The second of the measure is the Objective Edge Based measure (QE) that Xydeas and Petrovic proposed [7]. QE evaluates the amount of edge information that is transferred from source images to the fusion result. The edge preservation is connected to the texture information. The range of QE is between 0 and 1. It is considered a good measurement when it is close to 1. Q^{AF} and Q^{BF} are edge preservation values of input image A and B which is related fusion image F. w^A and w^B are weights determined by the ratios of the input image and fusion image edge strengths. N and M are sizes of the images.

$$
QE_p^{AB/F} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M} Q^{AF}(i,j)w^A(i,j) + Q^{BF}(i,j)w^B(i,j)}{\sum_{i=1}^{N}\sum_{j=1}^{M}\left(w^A(i,j) + w^B(i,j)\right)} \tag{7}
$$

We compared our proposed method to a conventional approximation image combining method [8].

Various pixel based image fusion algorithms use average or Burt's weight rule for approximation level combining rule to reduce computational costs. [2][4][5][6][9] and [10] are typical cases which use and introduce these rules. Fig. 5,6,7,8 are the result of each method. (c) is a result of the conventional method. (d) is a result of the proposed method. With human perception, background of images that is applied Burt's rule look cloudy because of background data of IR image. But images that applied the proposed method on (d) look clean. This is also observed on the evaluation test. Table 1 and 2 show each experimental result for GCF and QE. Clean image also means that image with proposed method have better contrast. Proposed method gets higher value of GCF then the other methods.

TABLE I. THE RESULT COMPARISON OF GCF

Methods	Evaluation result of each scenes			
	Night	***Day***	***Snowy***	***Lens flare***
Burt's	3.931	7.127	6.346	3.047
Proposed	4.036	7.141	6.523	3.068

TABLE II. THE RESULT COMPARISON OF QE

Methods	Evaluation result of each scenes			
	Night	***Day***	***Snowy***	***Lens flare***
Burt's	0.609	0.625	0.703	0.570
Proposed	0.622	0.666	0.742	0.587

Table. 2 also shows that our method can preserve input images' edge and orientation well. Edges of IR image also exist in visible image because IR image shows contour and shape of object on visible image too. This result tells us that the proposed method is not only good for human perception but also improve fusion performance.

IV. CONCLUSION

In this paper, we proposed the IR image transformation and the novel approximation image combining rule for a multi scale decomposition based image fusion. We use statistical data of a visible image for transformation to refine major information of IR image. Approximation image combining rule is made for keeping visible image and reducing lens flare effects at night. Based on the experimental results, GCF and QE, the proposed method provides good results both visually and quantitatively.

ACKNOWLEDGMENT

This research was supported by Seoul R&BD Program (WR080951)

REFERENCES

[1] S. Mallat, "Wavelet for a Vision," Proceeding of the IEEE: Special Issue on Wavelets, 84: 604-614, 1996.

[2] O. Rockinger,T. Fechner, "Pixel-Level Image Fusion: The Case of Image Sequences," SPIE Proceedings, 3374: 259-280, 1998.

[3] K. Matkovic, L. Neumann, A. Neumann, T. Psik, W. Purgathofer, "Global contrast factor—a new approach to image contrast, " Computational aesthetics in graphics, visualization and imaging, 159–68, 2005.

[4] Q. Miaw, B. Wang, "A Novel Image Fusion Method Using Contourlet Transform," International Conference on Communication, Circuits and System Proceedings, 548-552, 2006.

[5] P.Hill, N.Canagarajah, D.Bull, "Image Fusion using Complex Wavelets," 13th British Machine Vision Conference, 487–496, 2002.

[6] Z. Zhang, R. Blum, "A categorization of multi scale decomposition based image fusion schemes with a performance study for a digital camera application," Proceedings of the IEEE 87(8):1315–1326, 1999.

[7] V. Petrovic, C. Xydeas. “Objective image fusion performance measure,” Electronics Letters, 36(4):308–309, 2000.

[8] P. Burt, “A gradient pyramid basis for pattern selective image fusion,” Society for information Display, Digest of Technical Papers, 467-47, 1992

[9] Z. Xue, R. Blum, “Concealed weapon detection using color image fusion,” 6th International Conference of Information Fusion, 622-627, 2003.

[10] G. Piella, “A general framework for multi resolution image fusion: from pixels to regions,” Information Fusion 4(4):259-280, 2003.

(a) Visible image (b) LWIR image

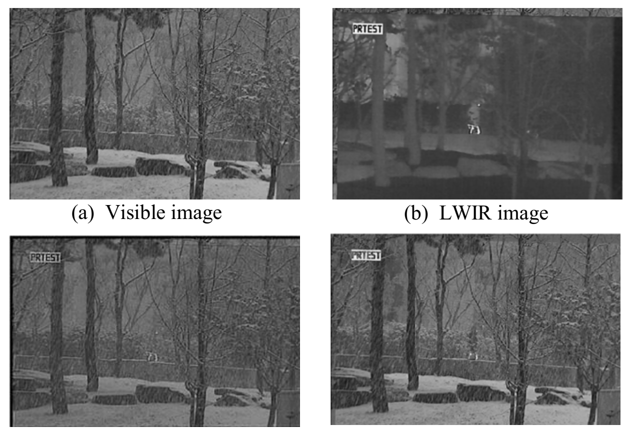
(c) Burt’s combining rule (d) Proposed combining rule

Figure 5. Result comparison : night scene

(a) Visible image (b) LWIR image

(c) Burt’s combining rule (d) Proposed combining rule

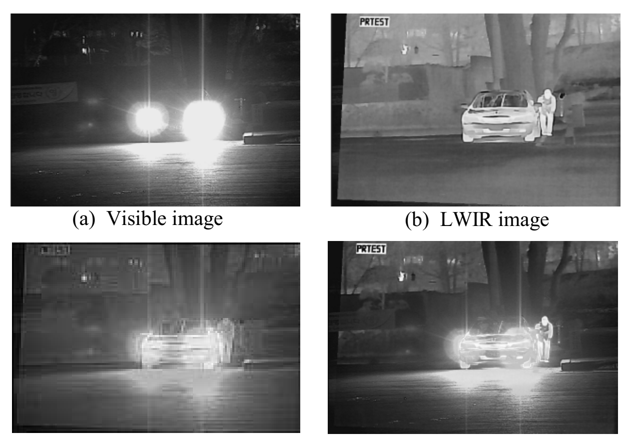
Figure 6. Result comparison : day scene



(a) Visible image (b) LWIR image

(c) Burt’s combining rule (d) Proposed combining rule

Figure 7. Result comparison : snowy scene



(a) Visible image (b) LWIR image

(c) Burt’s combining rule (d) Proposed combining rule

Figure 8. Result comparison : lens flare scene