A Hybrid Thermal-visible Fusion for Outdoor Human Detection

Ezrinda Mohd Zaihidee[1], Kamarul Hawari Ghazali[1], Jinchang Ren[2] and Mohd Zuki Salleh[3]
[1]*Faculty of Electrical & Electronics Engineering, Universiti Malaysia Pahang*
[2]*Department of Electronic and Electrical Engineering, University of Strathclyde*
[3]*Faculty of Industrial Sciences & Technology, Universiti Malaysia Pahang*
ezrindamz@gmail.com

***Abstract*—Multisensory image fusion can be used to improve the visual interpretability of an image for further processing task. A hybrid thermal-visible image fusion is proposed in this paper to detect the target and produce an output image that had all the information from both sensors. The thermal target region was extracted using Niblack algorithm and some morphological operators. Then, the source images were decomposed at pixel level using Stationary Wavelet Transform (SWT). The appropriate fusion rule were chosen for low and high frequency components and finally the fused image was obtained from inverse SWT. Results show that the proposed method achieve better to include figure of merit than the other three methods.**

***Index Terms*—multisensory image fusion; Niblack algorithm; Stationary Wavelet Transform; thermal-visible image fusion.**

I. INTRODUCTION

Image fusion systems are widely used in medical diagnosis, remote sensing, surveillance, navigation guidance and agriculture. Multisensory image fusion is a technique that combine two or more images that capture the same scene but with different imaging conditions (sensor types, weather conditions, time, polarization modes etc.) to obtain an image with highly comprehensive and quality for further processing tasks [1][2]. The concept of information fusion in military defense systems was proposed in a research report of the American military aspect in 1972 [3]. From the concept being proposed until today, numerous studies on new approaches and methods have been developed extensively to meet current demand.

Nowadays, there is a high demand in detecting and tracking human beings in surveillance systems. Among items monitored are abnormal behavior, human gait characterization and person identification [4]. These items required continuous monitor. There are many Closed Circuit Television (CCTV) being used for monitoring any suspicious activities but the cameras work well in controlled condition only and not able to work as a real time alarm system without human control [5]. Visible camera more accord to human visual characteristics and relatively high contrast ratio. In other words, it is not easy to make observation under low lighting conditions and in an uncontrolled outdoor environment [6]. To overcome this short comings, thermal camera was proposed for video surveillance system. It capture thermal radiation from different objects. Hence, it has obvious target information regardless of lighting conditions, haze or fog. However, it is difficult to identify the actual activity due to lack of details in the captured image. The low value of signal-to-noise ratio and contrast level [7] affect the quality of an image. Thus, combining both images will result in better awareness. Thermal-visible image fusion will provide better images as each image contains complementary information, different contrast level and resolution although they depict the same region at the same time [8].

In recent years, the research on thermal-visible image fusion pays particular attention to detect human at pixel level based on multiscale transforms (MST) where it consists three general steps. Firstly, decompose the input images into a multiscale transform domain. Then, merge the low and high frequency coefficients in each layer using appropriate fusion strategies. Finally, reconstruct the fused image at every layer to get the final image for further processing [9].

There are many techniques in image fusion using MST either in continuous or discrete domain. Curvelet transform in continuous domain is difficult to sample, therefore the edges of the final fused image are not smooth [10]. The contourlet transform in discrete domain [11][12] that inspired by curvelet transform produce a better result in terms of image smoothness. However, it is not translation invariant due to upsampling dan downsampling. Then, many researchers tend to introduce other techniques in thermal-visible fusion for human detection such as non-subsampled countourlet transform [13] [14], saliency extraction with multiscale decomposition [15][16], gradient transfer fusion [17], MST with sparse representation [18] and others.

This paper proposed a hybrid thermal-visible fusion for human detection based on thermal target extraction and Stationary Wavelet Transform (SWT). The target region in input thermal image was extracted using Niblack algorithm and some morphological operators. At the same time, both input images was decomposed into low and high frequency coefficient using SWT. Then, appropriate fusion rule was chosen to integrate the coefficients. Finally, a fused image was