Optical and Acoustical Methods in Science and Technology 2013

Hidden Object Detection System Based on Fusion of THz and VIS Images

M. Kowalski^{*}, N. Palka, M. Piszczek and M. Szustakowski

Institute of Optoelectronics, Military University of Technology, S. Kaliskiego 2, 00-908 Warsaw, Poland

Terahertz imaging makes it possible to acquire images of objects concealed underneath clothing by measuring the radiometric temperatures of different objects on a human subject. Finding objects hidden under clothes is one of key issues of public places security. A possible solution of this problem is a multispectral, terahertz-based screening system with implemented image fusion algorithm. The result of the fusion is a new image which gives more information and is more suitable for human perception. An imaging system is usually composed of various subsystems. Many of the imaging systems use imaging devices working in various spectral ranges. The aim of our studies is to detect and visualize objects hidden under clothing. Detection of potentially dangerous objects is important, but without proper visualization, can be difficult to utilize. Our goal is to build a system harmless to humans for screening and detection of hidden objects using a THz and VIS cameras. The presented system draws attention not only on the properties offered by a single camera but also on the potential of synergetic effects of the system by using the fusion of data registered by different sensor units. In this paper we present the multispectral passive imaging system for hidden threats visualization based on THz and VIS cameras. DOI: 10.12693/APhysPolA.124.490

PACS: 07.05.Pj, 42.30.Tz, 02.60.-x

1. Introduction

Terahertz radiations have shown unique properties and these radiations provide a variety of applications and opportunities in different fields. One of the most interesting applications of terahertz waves is imaging. The terahertz range of electromagnetic radiation has large potential in the field of hidden objects detection because it is not harmful to humans. However, the main difficulty in the THz imaging systems is low image quality due to low sensitivity and a small number of pixels in detecting modules of cameras.

Considering the fact that even THz images with low pixel resolution still provide valuable information, it is justified to combine them with the high-resolution images from a visible camera [1]. Image fusion is a process of combining two or more images, or it can be defined as the process by which several images or parts of images (features) are combined into a single image. This process allows us to combine the most interesting elements of images from various spectral ranges into one image.

A very important aspect of applying various processing techniques to images is proper assessment of the image quality. Performance measures are essential to determine the possible benefits of fusion as well as to compare the results obtained with different algorithms [2]. During our investigations many various algorithms were implemented and tested. Assessment of performance plays crucial role in the process of choosing the best image fusion method.

The aim of our studies is to detect and visualize objects hidden under clothing. Detection of potentially dangerous objects is important, but without proper visualization can be difficult to utilize. Security applications very often require a person to operate a system, thus the fused image is intended for presentation to a human observer for easier and enhanced interpretation. THz images are more difficult to assimilate for a human eye than visual images because they present a reality invisible for humans. Visible images are natural for human vision. The system is based on two cameras — passive THz and VIS. Our studies show that the fused image should contain information mostly from the visible image with superimposed elements from the THz image showing a detected object.

In this paper we present a hardware configuration of the imaging system, the results of various image fusion methods and propose a method of assessing quality of the fused image.

2. Imaging environment

Terahertz radiation is electromagnetic radiation in a frequency interval from 0.3 to 10 THz (1 mm–30 μ m wavelength). This band occupies a large portion of the electromagnetic spectrum between the infrared and microwave bands. Similar to other frequency bands like microwaves and infrared radiation, THz waves can be used as imaging media in a variety of applications.

Our goal is to detect objects hidden under clothes using images acquired by a THz camera. The object detection process means processing and fusion of images acquired by THz camera and a camera working in a visible spectrum of light. For our studies, the TS4 camera from ThruVision Systems Ltd. [3] was selected. The camera offers quite good quality images with the pixel resolution of 124×271 points and it is operating at 250 GHz. The camera working in the visible spectrum was manufactured by Axis and has a resolution of 600×550 pixels.

Because of the fact that testing our system requires a person carrying hidden object we used a thermal phan-

^{*}corresponding author; e-mail: mkowalski@wat.edu.pl

tom whose task is to imitate a human body (shapes as well as a temperature). Human phantom was placed on a moving platform and mimics a moving human being.

Another element necessary to build the experiment environment is hidden object. Objects imitating guns, knife or explosives were used as potentially dangerous objects hidden under clothing.

3. Fusion of THz and VIS images

The basis of system operation is the connection of images registered by the THz and VIS cameras into one image in order to obtain visualization understandable for human being. Image fusion is defined as the process of combining substantial information from several sensors using mathematical techniques in order to create a single composite image that will be more comprehensive and more useful for a human operator or other computer vision tasks. Current technology in imaging sensors offers a wide variety of information that can be extracted from an observed scene. Multi-sensor information is jointly combined to provide an enhanced representation in many cases of experimental sciences. In our study image fusion is a way to enhance the visibility of selected features (objects hidden under clothing) in the observed scene.

Our recent investigation on combining THz and VIS images shows that there is a great potential in combining two images from different spectral ranges especially for security applications. In this part of our research, we focus on extraction of the most valuable features from each image and combination them in one.

Image fusion methodology used in our investigations consists of three major steps. During the first step, the imaging material was acquired by two cameras — the THz and VIS. Since the THz and VIS images have different fields of view and sizes we found a common area for both images [4]. This operation was the second step of the fusion process and is very often connected with geometrical transformations of images. In the second step the VIS image was subjected to a cropping operation while the THz image was thresholded. During the third step, one of selected image fusion method was applied to both, pre-processed images. The methodology of fusion process is shown in Fig. 1.

The following list of image fusion methods was used: ratio pyramid, gradient, shift invariant DWT (SIDWT), morphological pyramid, contrast pyramid, laplacian pyramid, averaging, principal component analysis (PCA), minimum, maximum, discrete wavelet transform (DWT), and filter subtract decimate (FSD) pyramid [5].

One of the most important aspects of presented results is that all of these images were processed with methods which can be efficiently implemented in a computer system and work in a real-time mode. From many sets of methods the best and the most efficient algorithms were selected and finally implemented in the computer-based system. According to laboratory test results images can



Fig. 1. Methodology of fusing images.

be processed in real-time with the camera operating with 30 frames per second.

4. Quality assessment of fused images

The fusion algorithms may introduce distortions or artifacts into the processed images. So, image quality assessment (IQA) is essential to determine the possible benefits of fusion as well as to compare results obtained with different algorithms. Furthermore, IQA methods are necessary in order to obtain an optimal setting of parameters for a specific fusion algorithm. All of the IQA methods can be divided into two groups — reference and no-reference methods. The reference methods use a reference image to measure the quality, while the no-reference methods use only a resultant image to compute an image quality rate.

While the presence of a reference image or information regarding the reference simplifies the problem of quality assessment, the practical applications of such algorithms are limited in real-world scenarios where reference information is generally unavailable at nodes where quality computation is undertaken. The performance of any IQA method can be described by its correlation with human subjective judgments of quality, since the human is the "receiver" of the visual signal. Practical application of IQA methods requires that these algorithms compute a perceptual quality efficiently.

IQA in the case of a fused image is a complex task because after a process of image fusion one obtains an image that is an effect of a transformation of two images (THz and VIS) and it is not possible to find any reference image. A very popular method of evaluating the quality of fused images is calculation of values of the peak signal to noise ratio, mean squared error or average difference [6, 7]. These methods work by exploiting the pixel differences between images, the correlation between images and changes in the histogram, respectively. Our approach to assessing fused image quality is to evaluate perception of images by a human eye. This was the reason we used two kinds of IQA methods — the reference and the no--reference. We propose to use structural similarity index (SSIM) as the reference IQA method and the natural image quality evaluator (NIQE) as the no-reference (blind) quality assessment method. A combination of two unrelated image quality assessment methods can deliver the information required to assess the correctness of the fused image and select the most accurate fusion method.

The structural similarity index was invented for measuring the similarity between two images thus it is a full reference method [8]. Structural similarity is based on the idea that the human visual system is highly adapted to process structural information, and the algorithm attempts to measure the change in this information between reference and distorted image. According to numerous tests, SSIM is better at quantifying subjective image quality than mean squared error or peak signal to noise ratio [8, 9].

At a high level, SSIM attempts to measure the change in luminance, contrast, and structure in an image. The luminance is modeled as an average pixel intensity, contrast by the variance between the reference and distorted image, and structure by the cross-correlation between the two images. The resulting values are combined (using exponents referred to as alpha, beta, and gamma) and averaged to generate a final SSIM index value.

Structural information is the idea that the pixels have strong inter-dependences especially when they are spatially close [8]. These dependences carry important information about the structure of the objects in the visual scene. SSIM index values are in the range of $\langle 0, 1 \rangle$, where 1 is the value of two identical images. The SSIM index can be computed using the following equation [9]:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_x^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)},$$
(1)

where x, y are images, μ_x (μ_y) is the average of x (y), σ_x^2 (σ_y^2) are the variance of x (y), σ_{xy} is the covariance of x and $y, c_1 = (k_1L)^2, c_2 = (k_2L)^2$ are two variables to stabilize the division with weak denominator, L is the dynamic range of the pixel-values (typically $2^{\#\text{bits per pixel}} - 1$), $k_1 = 0.01$ and $k_2 = 0.03$.

The second IQA method proposed in this paper is NIQE. It is an IQA method that makes use of measurable deviations from statistical regularities observed in natural images without training on human-rated distorted images, and without any exposure to distorted images.

NIQE is based on the construction of a quality aware collection of statistical features based on a simple and successful space domain natural scene statistic model. These features are derived from a corpus of natural, undistorted images [10]. The quality of the distorted image can be described as the distance between the natural scene statistic feature model and the multivariate Gaussian fit to the features extracted from the distorted image and can be computed using following equation [10]:

 $D(v_1, v_2, \Sigma_1, \Sigma_2)$

$$= \sqrt{\left((v_1 - v_2)^{\mathrm{T}} \left(\frac{v_1 + v_2}{2} \right)^{-1} (v_1 - v_2) \right)}, \qquad (2)$$

where v_1 , v_2 are the mean vectors of the natural multivariate Gaussian model and the distorted image's multivariate Gaussian model, Σ_1 , Σ_2 are the covariance matrices of the natural multivariate Gaussian model and the distorted image's multivariate Gaussian model. NIQE values are in the range of $\langle 0, \infty \rangle$.

5. Results

The results of applying image fusion methods to sample images as well as the results of IQA of fused images are presented in Table.

Images presented below show that there is a huge potential for THz image processing methods. One of the most important aspects of presented results is that all of these images were processed with methods which can be efficiently implemented in a camera and work in a real-time mode. From many sets of methods the best and the most efficient algorithms were selected and finally implemented in the camera. According to laboratory test results images can be processed in real-time with the camera operating with 30 frames per second.

6. Summary

Despite the fact that terahertz images have low signal to noise ratio and low contrast, we have successfully achieved our goal to automatically detect and segment concealed objects. We build a system for hidden objects detection based on the passive THz camera and a VIS camera.

Based on assumptions made at the beginning, we compared many series of the IQA results to select the best image fusion method. According to our initial assumptions the value of the SSIM index of the VIS-fused images should tend to 1 while the value of SSIM index of the THz-fused images and the value of NIQE should tend to 0. This means that the correlation index between the fused and the VIS images is bigger than between the fused and the THz ones. It is natural because the VIS image is the background image for the fusion process, and only a few necessary and relevant elements of THz image are used. After analysis of many series of results the averaging method was chosen. The selected method was modified by replacing average with weighted average.

The most recent developments in THz technology and new explorations of THz imaging have advanced THz applications in the defense and security fields. There are still challenges and obstacles to implement THz technologies for defense and security in real-world applications. Powerful THz sources and sensitive THz detection methods need further development.

	Ratio pyramid	Gradient	SIDWT	Morphological pyramid	Contrast pyramid	Laplacian pyramid
NIQE	21.5012	12.0009	6.9006	8.2145	12.0555	8.3000
SSIM VIS-fused	0.1768	0.1817	0.1394	0.1750	0.0790	0.1400
SSIM THz-fused	0.0347	0.0711	0.1134	0.0517	0.0511	0.1399
	Averaging	PCA	Min	Max	DWT	FSD pyramid
			1			
NIQE	9.9978	33.1324	11.8390	16.4291	7.9581	11.8604
SSIM VIS-fused	0.2699	0.1230	0.1721	0.2205	0.1214	0.1779

Results of image fusion methods and quality evaluation.

References

- [1] V.A. Trofimov, Proc. SPIE 8189, (2011).
- [2] M. Piszczek, M. Kowalski, M. Szustakowski, Acta Phys. Pol. A 120, 720 (2011).
- [3] ThruVision System Ltd. website: http://www. thruvision.com.
- [4] C. Toxqui-Quitl, Proc. SPIE 5558, (2004).
- [5] M. Kowalski, M. Piszczek, N. Palka, M. Szustakowski, *Phot. Lett. Poland* 4, 3 (2012).
- [6] M. Sumathi, R. Barani, Proc. IEEE, PRIME 2012, p. 312.

- [7] W. Pei, G. Wang, X. Yu, *Proc. IEEE*, *IGARSS 2012*, p. 2280.
- [8] D. Brunet, E.R. Vrscay, Z. Wang, *IEEE Trans. Image Proc.* 21, 1488 (2012).
- [9] Z. Wang, A. Bovik, E.P. Simoncelli, *IEEE Trans. Image Proc.* 13, 4 (2004).
- [10] A. Mittal, R. Soundararajan, A.C. Bovik, *IEEE Sig. Proc. Lett.* **20**, 209 (2013).