

Original Research Paper

Metric for the Fusion of Synthetic and Real Imagery from Multimodal Sensors

¹Rami Nahas and ²S.P. Kozaitis

¹Electrical Engineering, University of Business and Technology, Jeddah, Saudi Arabia

² Department of ECE, Florida Institute of Technology, Melbourne, USA

Article history

Received: 20-11-2014

Revised: 29-11-2014

Accepted: 18-12-2014

Corresponding Author:

S.P. Kozaitis

Department of ECE, Florida

Institute of Technology,

Melbourne, USA

Email: kozaitis@fit.edu

Abstract: We described a method to improve night vision for vehicle navigation that combined images from a thermal infrared camera and from a public database of stored images. Such an approach allows a night scene to appear as if it were daytime for automotive applications thereby increasing safety. We described a new metric to evaluate the fusion of such an augmented reality system and compared leading fusion algorithms to determine the efficacy of our approach.

Keywords: Image Fusion, Night Vision, Augmented Reality, Thermal Imagery

Introduction

Darkness and poor visibility at night limit what a vehicle operator can detect causing an unsafe navigation condition. Ideally, a system is needed that can make low-visibility situations like driving at night appear as if they were in daytime, as such a system would provide a particularly effective approach for vehicle navigation in darkness.

There already exist a variety of systems to aid vehicle navigation in low-visibility situations. Thermal imagers have been reported to identify pedestrians or animals at night (Li *et al.*, 2010; Geronimo *et al.*, 2010, George *et al.*, 2012). Warm objects can be displayed on a supplemental screen and even segmented and identified. Augmented systems can be used to help identify lane markings where visibility is limited. However, although these systems can be useful, they focus on specific situations and a more general approach is needed that can be used for a variety of conditions.

Infrared or thermal cameras have often been used in night vision systems, but they provide limited data about a scene. Combining images from additional sensors can improve results. A range of night vision systems have been developed to improve the ability to see at night (Bishop, 2000; Vu *et al.*, 2012; Bhatnagar *et al.*, 2011). However, problems still remain. Thermal infrared cameras are not particularly good at capturing some elements of a scene such as trees, leaves and grass in a natural landscape. These are not visible at night with visible-light sensors, so adding additional sensors has been one traditional approach to this problem.

In contrast to these subsequent approaches, we used stored images of a desired scene acquired in daytime to augment the same scene acquired at night in real-time. The public database Google Earth is a ready asset that is being continuously updated by multiple sources. It contains daytime images of a large number of scenes from around the world. Rather than use additional sensors, our approach combines sensor data acquired from a thermal camera with the stored daytime image from Google Earth. In this way, we augment a real-time scene with the image of that scene from the database. Such an idea has been previously proposed but involved color mapping rather than image fusion (Qadir *et al.*, 2014). Because the database does not provide real-time information, we prioritized the fusion of information between the real-time and database data using a metric. In the next sections, we describe our system followed by a description of our metric. We then compare two different scenes with and without a pedestrian with different fusion algorithms.

System

As previously stated, we created a system that uses a stored database of visible daytime images to enhance thermal imagery captured at night. Basically, information from the stored database of images would be used when there is little or no thermal signal. This could improve safety by identifying objects or features that are not visible to the naked eye or thermal camera.

The system inputs are from the thermal camera and a GSP sensor; the visible display is the output as illustrated in a block diagram of the system shown in Fig. 1.

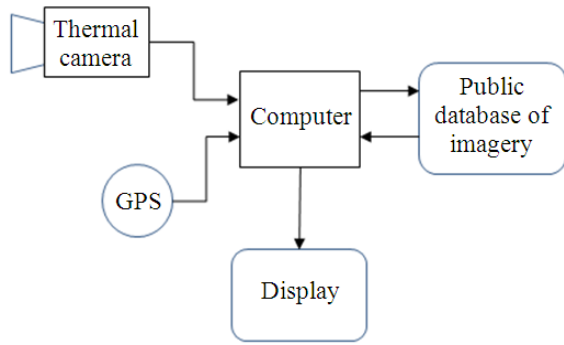


Fig. 1. Block diagram of system

The camera's input is that of a night scene and the GPS sensor provides the vehicle coordinates. These inputs constitute two of the three inputs to the computer. Based on the GPS coordinates, a wireless request is sent to a public database that also includes camera angle information and it returns a visible daytime image of the scene corresponding to the thermal camera image (Salmen *et al.*, 2012). The output of the system is the fusion of the registered thermal and database images. The process is then repeated based on the difference in location of the previously requested image.

Metric Selection

Most image fusion metrics were designed to consider imagery from sources of the same scene taken at the same time. Our application demands that one image has priority over the other because the thermal image contains live data whereas the database image is used to enhance the thermal imagery. Although image fusion metrics can be adjusted to weight one image when compared to another, we introduce a ratio that gives priority of the thermal image over the visible image. We show that our ratio is more definitive in terms of information than a weighting factor and illustrate the significance of the metric with two examples.

We initially computed the Mutual Information (MI) between the each source image where MI_X and MI_Y refer to the mutual information of the thermal and database images respectively. We can write our metric as:

$$MI_{total} = MI_X + MI_Y \{MI_X / MI_Y = b\} \quad (1)$$

where, β is a fixed value. The metric uses a constant ratio of the two values of mutual information in forming the sum of two values.

We used a simple fusion algorithm to examine the effect of the ratio β . In our particular case, the thermal image is represented by X and the database image by Y . The fusion algorithm uses pixels values from either one image or the other depending on a threshold. The fused image was represented by F and described as Equation 2:

$$F = X \text{ if } X > T, \text{ or } F = Y, \text{ otherwise,} \quad (2)$$

on a pixel-by-pixel basis, where T is a threshold.

Results

We used two different scenes to gather results and acquired thermal imagery with and then without a pedestrian present in each case. The database images were acquired from the Google Street View database. We compared three different fusion algorithms, as well as the simple one algorithm described in Equation 1. The additional fusion algorithms were based upon a Laplacian Pyramid (LAP), Shift Invariant Discrete Wavelet Transform (SIDWT) and Principal Component Analysis (PCA) methods. Initially, the two images were combined into an intermediate result R using a fusion algorithm. Then, the final result was described as Equation 3:

$$F = X \text{ if } X > T_1, F = Y \text{ if } X < T_2, \text{ or } F = R, \text{ otherwise,} \quad (3)$$

where, T_1 and T_2 are thresholds. The idea here is that the data is combined when data from both sensors are similar. Therefore, the results from the fused image R were applied when pixel values of the input images were between two thresholds. The thermal and database images are shown in Fig. 2.

We examined results for $\beta \sim 0.2, 0.9$ and 1.5 . These values corresponded to when the database image had priority, the two source images had about equal priority and when the thermal image had priority, respectively. To obtain these values of β , the threshold T was simply changed until that desired value was obtained.

Database Image Dominant

A value of $\beta \sim 0.2$ meant that the MI between the resulting fused image and the database image is a factor of five greater than the MI between the thermal and fused images. For this value of β , the fused image will most likely be dominated by the visible image except for in relatively warm areas of the thermal image. Figure 3 and 4 showed the results of the four different algorithms for the two different scenes. Table 1 and 3 show the values of the MI for these images, the total MI and the MI of the thermal and database images written as MI_X and MI_Y respectively.

Both Fig. 2 and 3 show similar results in that the pedestrian is clearly visible with most of the remaining image due to the database image. Each term of the MI as well as the total value was highest for the simple algorithm. Table 2 and 4 show the MI values of the same scene but without the pedestrian present.

The results are similar to that of when the pedestrian was present. However, the MI between the thermal and fused images is generally lower in the scene without the pedestrian. This makes sense because the pedestrian comes entirely from the thermal image.



Fig. 2. Images used in experiments (a) thermal image of scene 1 with pedestrian (b) thermal image of scene 2 with pedestrian (c) visible image of scene 1 from database (d) visible image of scene 2 from database

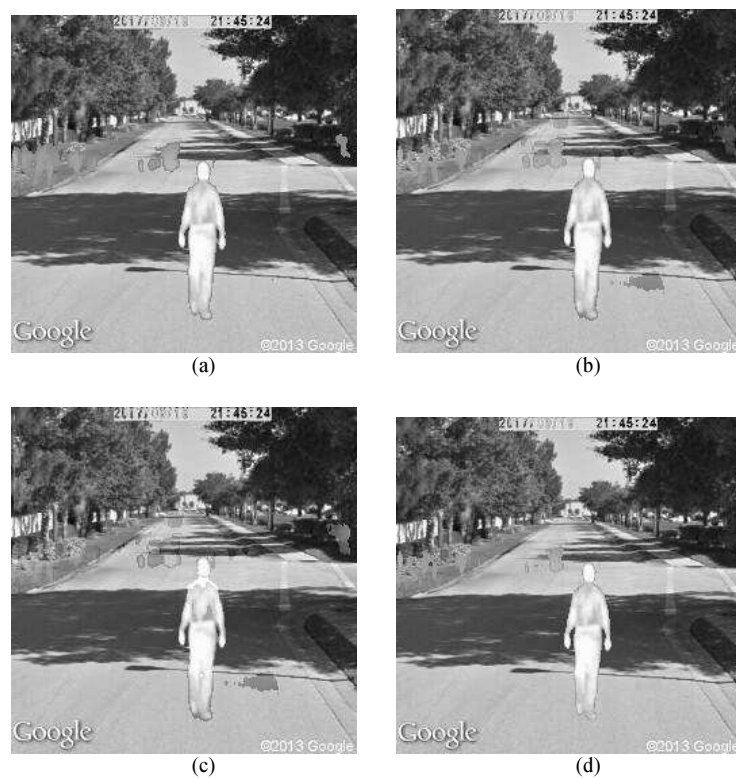


Fig. 3. Fusion results of scene 1 using $\beta \sim 0.2$ for different algorithms (a) simple (b) LAP (c) SIDWT (d) PCA

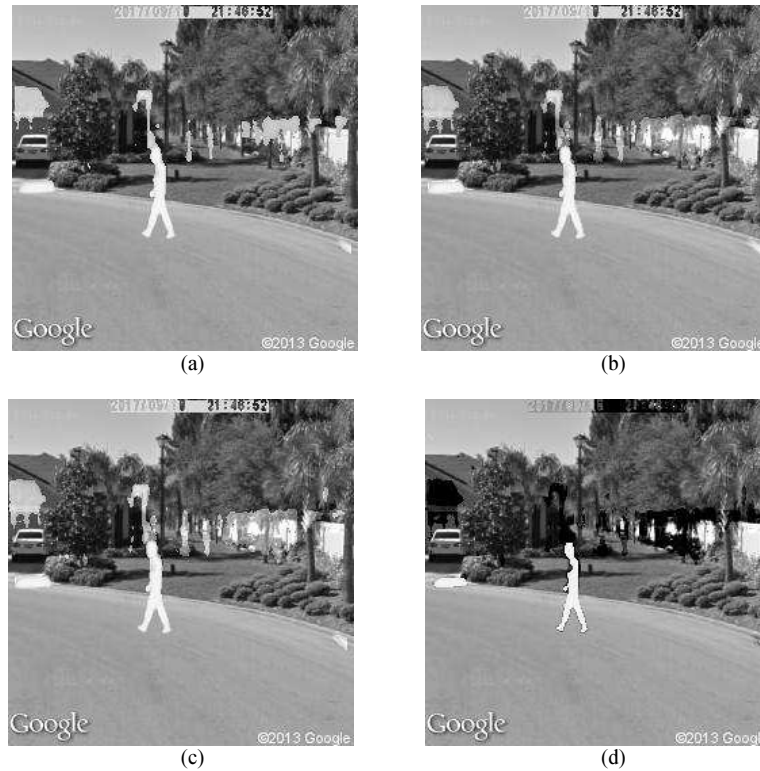


Fig. 4. Fusion results of scene 2 using $\beta \sim 0.2$ for different algorithms (a) simple (b) LAP (c) SIDWT (d) PCA

Table 1. Mutual information results of scene 1 with $\beta \sim 0.2$

	Simple	LAP	SIDWT	PCA
MI_{total}	0.538	0.514	0.513	0.534
MI_X	0.090	0.087	0.087	0.089
MI_Y	0.448	0.427	0.426	0.445
β	0.201	0.204	0.204	0.200

Table 2. Mutual information results of scene 1 without pedestrian present $\beta \sim 0.2$

	Simple	LAP	SIDWT	PCA
MI_{total}	0.527	0.439	0.454	0.432
MI_X	0.090	0.075	0.076	0.074
MI_Y	0.437	0.364	0.378	0.358
β	0.206	0.206	0.201	0.207

Table 3. Mutual information results of scene 2 with $\beta \sim 0.2$

	Simple	LAP	SIDWT	PCA
MI_{total}	0.568	0.539	0.540	0.521
MI_X	0.095	0.091	0.092	0.089
MI_Y	0.473	0.448	0.448	0.432
β	0.201	0.203	0.205	0.206

Table 4. Mutual information results of scene 2 without pedestrian present $\beta \sim 0.2$

	Simple	LAP	SIDWT	PCA
MI_{total}	0.569	0.436	0.478	0.510
MI_X	0.096	0.073	0.080	0.087
MI_Y	0.473	0.363	0.398	0.423
β	0.203	0.201	0.201	0.206

Images with Similar Priority

A value of $\beta \sim 0.9$ meant that the MI between the resulting fused image and the database image was similar to that of the MI between the thermal and fused image. For this ratio, the visible image and thermal image will have a similar amount of information in the fused image. Figure 5 and 6 show the results of the four different algorithms for the two different scenes. Table 5 and 7 shows the values of the MI for the two scenes with the pedestrian and Table 6 and 8 show the results without the pedestrian.

Figure 5 shows that the simple and PCA results look similar but different from that of the other two algorithms, LAP and SIDWT, which look similar to each other. The results from the simple and PCA algorithms visually contained more data from the thermal image. The MI values were also higher in these cases when compared to the other two algorithms. However, the MI values of the simple algorithm were higher than in the PCA case. Figure 6 gave similar MI results as that of Fig. 5, but the image using the PCA algorithm seemed to have more contribution from the visible image in areas of vegetation, with less in the road region. When the pedestrian was not present, the MI of the simple algorithm remained similar, but the MI values from the PCA result dropped significantly.

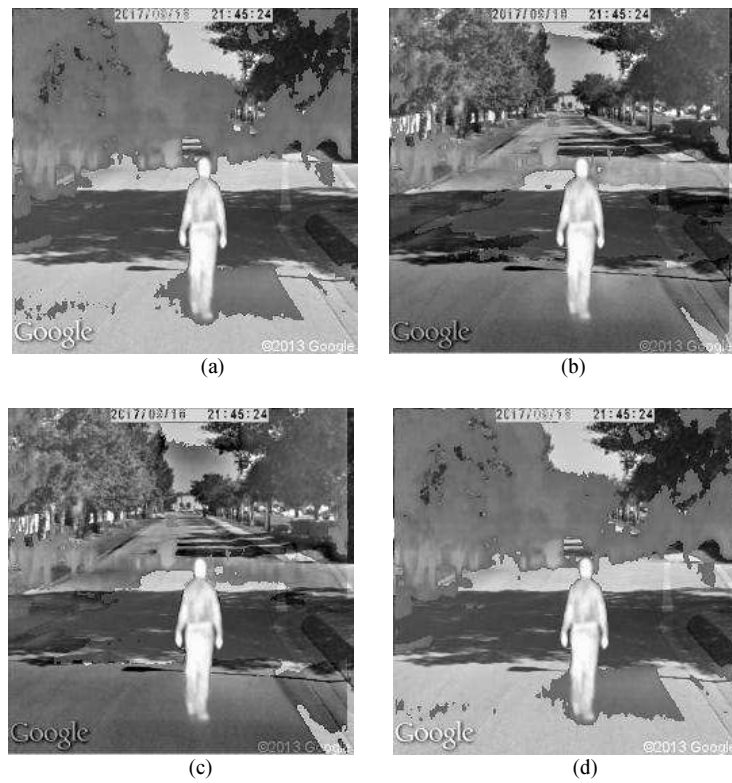


Fig. 5. Fusion results of scene 1 using $\beta \sim 0.9$ for different algorithms (a) simple (b) LAP (c) SIDWT (d) PCA

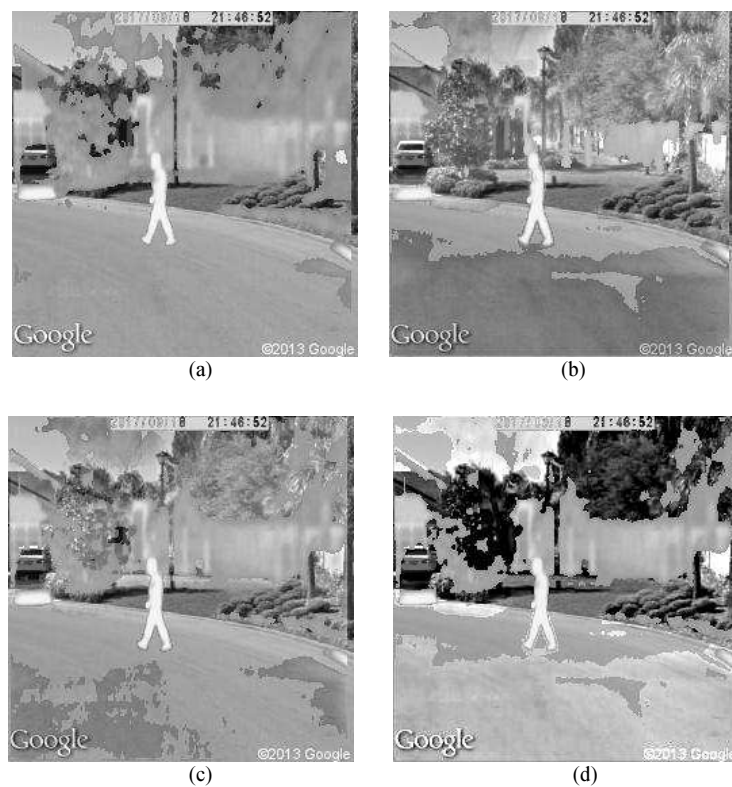


Fig. 6. Fusion results of scene 2 using $\beta \sim 0.9$ for different algorithms (a) simple (b) LAP (c) SIDWT (d) PCA

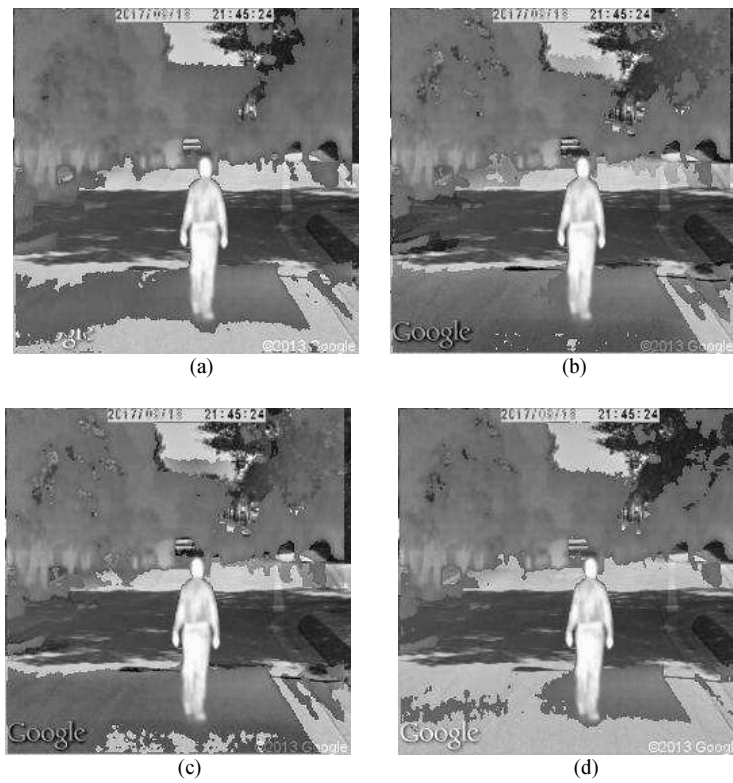


Fig. 7. Fusion results of scene 1 using $\beta \sim 1.5$ for different algorithms (a) simple (b) LAP (c) SIDWT (d) PCA

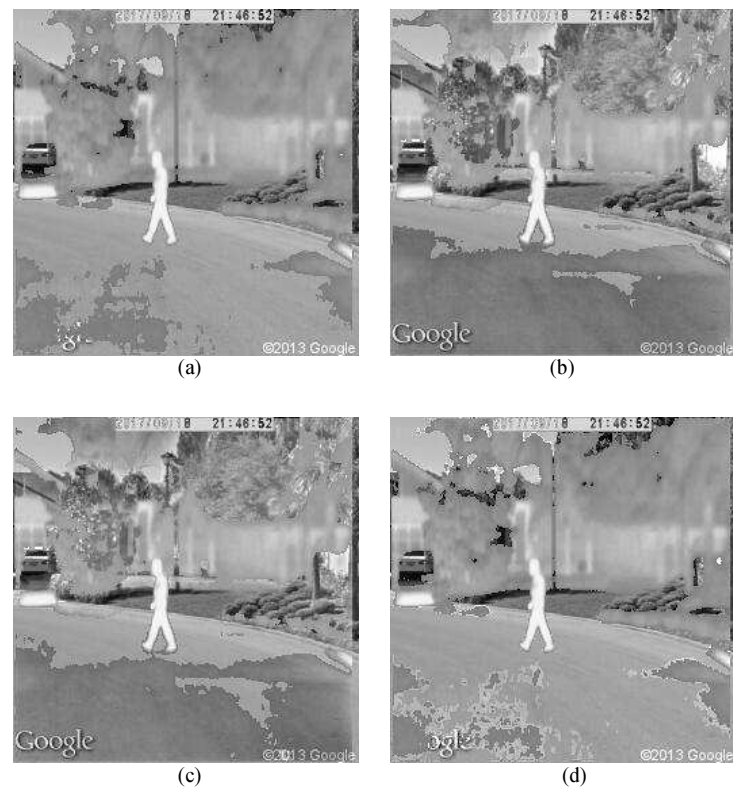


Fig. 8. Fusion results of scene 2 using $\beta \sim 1.5$ for different algorithms (a) simple (b) LAP (c) SIDWT (d) PCA

Table 5. Mutual information results of scene 1 with $\beta \sim 0.9$

	Simple	LAP	SIDWT	PCA
MI _{total}	0.533	0.221	0.226	0.528
MI _X	0.253	0.105	0.107	0.250
MI _Y	0.280	0.116	0.119	0.278
β	0.904	0.905	0.899	0.899

Table 6. Mutual information results of scene 1 without pedestrian present with $\beta \sim 0.9$

	Simple	LAP	SIDWT	PCA
MI _{total}	0.506	0.247	0.348	0.342
MI _X	0.240	0.117	0.165	0.162
MI _Y	0.266	0.130	0.183	0.180
β	0.902	0.900	0.902	0.900

Table 7. Mutual information results of scene 2 with $\beta \sim 0.9$

	Simple	LAP	SIDWT	PCA
MI _{total}	0.468	0.430	0.430	0.430
MI _X	0.222	0.204	0.204	0.204
MI _Y	0.246	0.226	0.226	0.226
β	0.902	0.903	0.903	0.903

Table 8. Mutual information results of scene 2 without pedestrian present with $\beta \sim 0.9$

	Simple	LAP	SIDWT	PCA
MI _{total}	0.498	0.270	0.274	0.428
MI _X	0.232	0.128	0.130	0.203
MI _Y	0.257	0.142	0.144	0.225
β	0.903	0.901	0.903	0.902

Table 9. Mutual information results of scene 1 with $\beta \sim 1.5$

	Simple	LAP	SIDWT	PCA
MI _{total}	0.523	0.381	0.416	0.463
MI _X	0.315	0.230	0.250	0.278
MI _Y	0.208	0.151	0.166	0.185
β	1.510	1.520	1.510	1.500

Table 10. Mutual information results of scene 2 without pedestrian present with $\beta \sim 1.5$

	Simple	LAP	SIDWT	PCA
MI _{total}	0.478	0.326	0.341	0.415
MI _X	0.287	0.196	0.205	0.249
MI _Y	0.191	0.130	0.136	0.166
β	1.500	1.510	1.510	1.500

Table 11. Mutual information results of scene 2 with $\beta \sim 1.5$

	Simple	LAP	SIDWT	PCA
MI _{total}	0.498	0.300	0.300	0.458
MI _X	0.299	0.180	0.180	0.275
MI _Y	0.199	0.120	0.120	0.183
β	1.500	1.500	1.500	1.500

Table 12. Mutual information results of scene 2 without pedestrian present with $\beta \sim 1.5$

	Simple	LAP	SIDWT	PCA
MI _{total}	0.486	0.380	0.228	0.333
MI _X	0.292	0.229	0.137	0.200
MI _Y	0.194	0.151	0.091	0.133
β	1.510	1.520	1.510	1.500

Thermal Image Dominant

A value of $\beta \sim 1.5$ indicates that the MI between the resulting fused image and the thermal image is a factor of 1.5 greater than the MI between the database and fused images. For this value of β , the thermal image should have a significant contribution to the final image. Figure 7 and 8 show the results of the four different algorithms for two different scenes. Table 9 and 11 show the values of the MI for these images containing a pedestrian and Table 10 and 12 show the results without the pedestrian.

Both Fig. 7 and 8 show similar results to when the images had similar priority. The results looked similar for the four algorithms, with the simple and PCA results appearing similar to each other and the remaining two algorithms appearing similar to each other. The MI values were also higher in the simple and PCA algorithm cases when compared to the other two. However, the MI values of the simple algorithm were higher than in the PCA case, much like the previous case.

Conclusion

We described an image fusion system that can lead to improved safety for nighttime vehicle operation. The use of a public database gives results not possible with other approaches. We introduced a new metric to prioritize images that was suited to our approach. We found that using the maximum of the sum of MI values while maintaining a constant ratio allowed us to compare fusion algorithms. We found that when giving the database image priority over the thermal image, the fused image appeared the most daytime-like and easily showed a pedestrian. As the priority of the thermal image increased, the background of the image was formed by both thermal and database images. A simple background replacement algorithm gave the best results in our experiments suggesting that one or the other image dominated a particular region. Developing improved fusion algorithms to properly combine such images could further improve the results for nighttime navigation.

Acknowledgement

Support of R. Nahas by the SACM is acknowledged.

Funding Information

Sponsorship of R. Nahas by the SACM is acknowledged.

Author's Contributions

R. Nahas: Participated in collecting and analyzing data and contributed to writing of the manuscript.

S. Kozaitis: Organized the study, participated in analyzing data and contributed to writing of the manuscript.

Ethics

Parts of this work is from R. Nahas's Ph.D thesis and has not been published elsewhere.

References

- Bhatnagar, G., Q. Wu and B. Raman, 2011. Navigation and surveillance using night vision and image fusion. Proceedings of the IEEE Symposium on Industrial Electronics and Applications, Sep. 25-28, IEEE Xplore Press, Langkawi, pp: 342-347. DOI: 10.1109/ISIEA.2011.6108728
- Bishop, R., 2000. Intelligent vehicle applications worldwide. IEEE Intell. Syst. Applic., 15: 78-81. DOI: 10.1109/5254.820333
- George, P., I. Thouvenin, V. Fremont and V. Cherfaoui, 2012. DAARIA: Driver assistance by augmented reality for intelligent automobile. Proceedings of the IEEE Intelligent Vehicles Symposium, Jun. 3-7, IEEE Xplore Press, Spain, pp: 1043-1048. DOI: 10.1109/IVS.2012.6232220
- Geronimo, D., A.M. Lopez, A.D. Sappa and T. Graf, 2010. Survey of pedestrian detection for advanced driver assistance systems. IEEE Trans. Patt. Anal. Mach. Intell., 32: 1239-1258. DOI: 10.1109/TPAMI.2009.122
- Li, J., W. Gong, W. Li and X. Liu, 2010. Robust pedestrian detection in thermal infrared imagery using the wavelet transform. Infrared Phys. Technol., 53: 267-273. DOI: 10.1016/j.infrared.2010.03.005
- Qadir, H., S.P. Kozaitis and E.A. Ali. 2014. Enhancing thermal video using a public database of images. Proceedings of the SPIE 9121, Multisensor, Multisource Information Fusion: Architectures, Algorithms and Applications, May 5-9, SPIE, USA. DOI: 10.1117/12.2050399
- Salmen, J, S. Houben and M. Schilipsing, 2012. Google street view images support the development of vision-based driver assistance systems. Proceedings of the IEEE Intelligent Vehicles Symposium, Jun. 3-7, IEEE Xplore Press, Alcala de Henares, pp: 891-895. DOI: 10.1109/IVS.2012.6232195
- Vu, A., A. Ramanandan, A. Chen, J.A. Farrell and M. Barth, 2012. Real-time computer vision/DGPS-aided inertial navigation system for lane-level vehicle navigation. IEEE Trans. Intell. Transport. Syst., 13: 899-913. DOI: 10.1109/TITS.2012.2187641