ISSN: 2976-7466

Journal of Biosensors and Bioelectronics Research

Research Article

Open Access

Using Color Feature Generation to Fuse Thermal and Reflective Images

Richard Vollmerhausen

RHV Electro-Optics, LLC Lake Mary, Florida, USA

ABSTRACT

This paper presents an image fusion method based on the observation that the human visual system integrates high-resolution spatial features and a low-resolution color map to form the perception of high-resolution color. Image fusion is implemented by letting the eye “color” a high-resolution near-infrared image with the features of a long-wave thermal infrared image. Several example uses of this fusion method are provided, in which image-intensifier and thermal images are fused to display one image with the prominent features of both constituent images. Other uses of this fusion method are also explored and examples provided.

***Corresponding author**

Richard Vollmerhausen, RHV Electro-Optics, LLC Lake Mary, Florida, USA.

Received: June 24, 2026; **Accepted:** June 29, 2026; **Published:** July 08, 2026

Keywords: Image Fusion, Feature Extraction, Multispectral Fusion, Color Vision, Thermal and Reflective Fusion, Color Coding, Digital Image Enhancement

Introduction

Multiplicative Fusion (MF) differs from other image fusion methods because the human vision system discovers and combines the saliant features of the constituent images. As a result, image features are combined in a way that supports cognitive interpretation. Further, MF produces artifact-free imagery, meaning that features appear in the fused output only if they are present in at least one of the source images. Also, MF is computationally simple to implement. The constituent images are registered and then pixel-by-pixel (array) multiplied.

MF is based on the observation that human color vision mechanisms can integrate the spatial features of two images, provided that they are registered views of the same scene. Digital image processing can take advantage of the vision mechanisms that add color features to the achromatic spatial features in the human visual cortex. Gouras’s Webvision articles on color vision outlines the visual system features leveraged by MF [1,2]. Additional details were provided by Lee and the practical applications described in this paper are consistent with Land’s demonstrations and Retinex color prediction algorithm [3,4].

We argue that empirical data describing color vision can be effectively utilized without requiring a detailed understanding of the physiological mechanisms underlying those data. The empirical data support the assertion that our vision system will combine the feature sets of thermal and reflective images of the same scene. Additionally, we created fused images to illustrate behavior and advantages of MF.

Presenting only a few fused images does not demonstrate the robustness of a fusion algorithm. However, no empirically validated objective metric that quantifies the efficacy of an image fusion method is available. Examples do not prove the general case, but neither does calculating metrics unsupported by empirical evidence.

In this paper, as an alternative to objective metrics, we describe empirical data supporting the premise that human vision combines the features of separate images. Because a multiplication in the spatial domain is a convolution in the frequency domain, the information in both images is represented in some fashion in the fused results. The open question is whether the salient features of the constituent images are visually interpretable in the fused results.

The remainder of this paper is structured as follows. Section 2 explains why we believe that our color vision mechanisms combine the features of multiple images in a cognitively understandable visual context, and Section 3 provides supporting examples. Conclusions are presented in Section 4.

Some Empirically Reported Characteristics of Color Vision

This section describes the use of experimentally determined features of human color vision to fuse thermal and reflective images.

Figure 1 depicts the layout of cone cells in the retinas of five individuals with normal color vision [5]. The cone types were identified using laser return measurements; gray areas indicate regions in which the laser signals could not be reliably interpreted. S, M, and L cones are shown in blue, green, and red, respectively. The blue cones are sparsely distributed, whereas the red and green cones vary in both quantity and spatial distribution, typically occurring in clusters.

Citation: Richard Vollmerhausen (2026) Using Color Feature Generation to Fuse Thermal and Reflective Images. Journal of Biosensors and Bioelectronics Research. SRC/JBBER-161. DOI: doi.org/10.47363/JBBER/2026(4)147

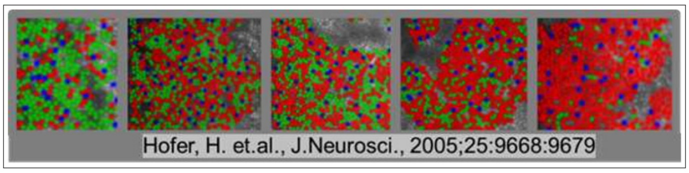

Figure 1: Maps of Cone types in the Retinas of Five Individuals with Good Color Vision [5]

Figure 2 illustrates how light from a scene, focused by a lens, forms an image on the retina. The retina is not designed to detect the color of light at every point in the scene; its spectral sensing resolution is inherently low, and directly generating a high-resolution color image from the retina does not appear to be feasible.

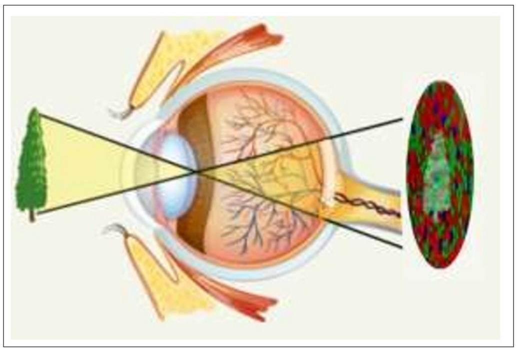

Figure 2: Light from the Scene is focused on the Retina where the S Cones are Sparse and the M and L Cones are Arranged in Clusters. The S, M and L Cones are not Arranged in Triads

Experimental findings from vision science indicate that the M and L cones jointly contribute to the formation of an achromatic feature set in the visual cortex, which is essential for generating the neural signals required for vision [1,2]. Gouras also considered the clumped spatial distribution of the M and L cones, as shown in Figure 1. Cone density in the retina varies significantly, ranging from approximately 200,000 cones per square millimeter at the fovea (on-axis) to approximately 30,000 cones per square millimeter two degrees off-axis. Despite this variation, even at the highest density, vision would appear patchy and mottled if the M and L cones did not jointly contribute to spatial resolution. Evidence indicates that the M and L cones and possibly the S cones as well coordinate to form a unified achromatic feature set, supporting the spatial acuity of human vision [1,2].

In summary, based on published experimental data, the M and L cones contribute together to an achromatic feature map. Our visual acuity strongly indicates that the M and L cones are uniformity-corrected to form the achromatic feature set.

According to Gouras, on the rare occasion that an individual has lost the functioning of the color region of the visual cortex, but the achromatic feature set remains healthy, that individual has full-resolution achromatic vision. Given that both M and L cones are known to contribute to achromatic vision, we postulate that both an achromatic image representation (a set of spatial features) and signals representing the strength of non-uniformity corrections exist in the vision system in some form. It is known that the visual cortex maintains a low-resolution color map and a single high-resolution luminance map. Some portion of the brain integrates low-resolution chromatic information with high-resolution spatial features, resulting in full-color, high-resolution perceptual images.

Figure 3 illustrates the proposed fusion concept. Two images (A and B) are derived from a high-resolution color image: grayscale and blurred color versions of the original. One image is made semi-transparent and the two are then combined via pixel-by-pixel multiplication. The result is a full-color, high-resolution representation of the original image (shown in the center of the figure).

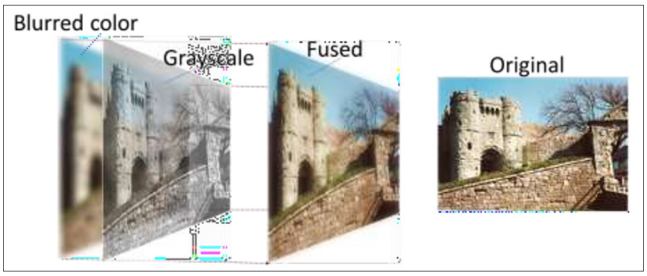

Figure 3: Blurred Color and Grayscale Images on the Left are Derived from the Original Image on the Right. The Fused Image is Obtained using Eq. 1

By making one image semi-transparent and performing pixel-by-pixel multiplication, a fused image is produced. Eq. 1 uses MATLAB code, where (.*) denotes array multiplication.

```
a = mean(B(:));
B = B + a
for k = 1:3
C(:, :, k) = A(:, :, k).*B(:, :)            (1)
end
a = min(C(:)) C = C – a
```

The physiological basis of human vision remains incompletely understood, even among vision scientists, and this study does not attempt to elucidate its mechanisms. Eq. 1 does not represent color vision behavior; instead, it presents a convenient digital processing method that imposes color feature patterns on the amplitude of high-resolution spatial features.

If the cones on the retina are non-uniformity-corrected over an area, then the gain-correction signals color the high-resolution features. In other words, the corrected cone matrix senses high-resolution spatial features, and the non-uniformity correction signals over some area of the retina represent the color difference between the cone types. Regardless of the inner workings of color vision, it is sufficient for our purposes that features derived from the retina convey both non-uniformity between cone types and luminance amplitude.

The multiplication in Eq. 1 mimics intensity variation due to color. Because of the non-uniformity correction, the high-resolution achromatic picture features are not corrupted. However, in a way unknown to us, that same non-uniformity correction causes our vision system to demarcate those spatial features by color.

The image fusion process involves replacing the low-resolution color image with a thermal image
T. The result is a monochrome fused image D. a = mean(B(:));

```
B = B + a
D(:, :) = T(:, :).*B(:, :)        (2)
a = min(D(:)) D = D – a
```

The reflective image is colored by the thermal image. However, the output image D need not be colored. Instead, D is equivalent to a grayscale conversion of a color image. If the thermal image is not low resolution, we do not know which picture, thermal or reflective, is coloring the other; however, if the features of both are visible, that is irrelevant.

The code snippet given as Eq. 2 represents the digital processing method used to fuse the examples shown in Section 3.

Examples of Multiplicative Fusion

The images shown in Figures 6, 7, 9, 10 and 11 were captured using the two cameras shown in Figure 4. The reflective camera was a PVS-14 image intensifier viewed using an HDMI CMOS video camera. The uncooled thermal camera used an HDMI video format. However, the tactical vehicle reflective imagery shown in Figure 5 was collected using an HDMI video camera, converted to grayscale, and bicubically interpolated to 590 × 401 pixels. Several frames generated by a 590 × 401-pixel long-wave thermal serial scanner were integrated to reduce noise in the tactical vehicle thermal images.

Figure 4: Reflective and Thermal Cameras used to Collect Most of the Fused Imagery

Figures 5A and B show thermal and reflective (visible) images, respectively, of a self-propelled Howitzer. The two images were captured at separate times and somewhat different aspects, so the registration is a few pixels off. The shapes of the turret, hatch, fenders, gun barrel, sides, wheels, and fenders are the same in both picture types. However, the shading of the top and edges of the front fender differ, and the visible picture shows the details of the front fender and vehicle sides. Additionally, the thermal image shows the vehicle heating from the engine and the usage of the road wheels. Figure 5C shows the features of each constituent image incorporated into a fused image using MF.

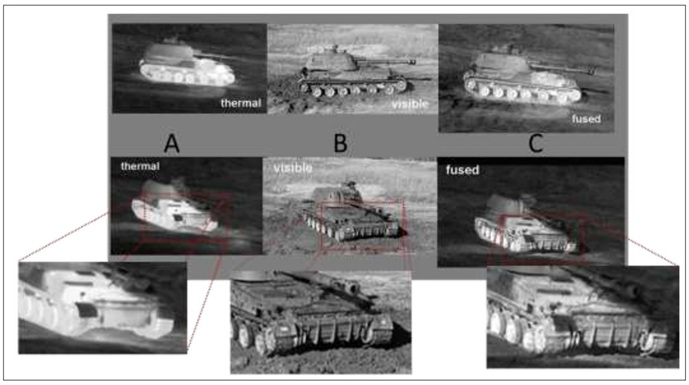

Figure 5: Thermal And Visible Images in A and B Were MF Fused as Shown In C

Figure 6 presents an example in which a kneeling person exhibits a strong thermal signature while the surrounding room remains cold, thus providing poor thermal contrast. Although the room lights were turned off, ambient stray light enabled the image intensifier to capture a clear image of the room and its contents.

Note that this example was the first fused imagery taken with the cameras shown in Fig. 4, and one video recording board was used for both cameras. Only in this first example, video capture was not simultaneous. The person in the pictures did not kneel in the same place after moving to switch video inputs.

Figure 6: A Kneeling Person is Highlighted in the Thermal Image and Room Contents are Visible in the Reflected Image. The Fused Image Shows Both the Person and Room Contents

Figure 7 shows a view from a family room into the living room of a house, where a woman is standing behind a chair. The thermal image reveals only the woman and a robot vacuum on the floor. In contrast, the image intensifier (I2) captures the details of the furniture and wall hangings, but the woman and the wall behind her are washed out by sunlight from the windows (not shown). The fused image successfully combines the strengths of both modalities, showing the details captured by the intensifier as well as the woman and the wall behind her.

Notably, the house located in Florida has metal studs in its walls. Consequently, faint dark stripes corresponding to the thermal signatures of the studs appear behind the woman in the fused image. Interestingly, these studs are not visible in either the standalone thermal or image-intensifier images.

Figure 7: The Fused Image Reveals Details That are not Visible in Either the Thermal or Reflective Image Alone. Note The Metal Studs Behind Wallboard

In Figure 8, a person is walking along a fence. The man is visible in the thermal image (A) but the fence is not, and the fence is visible in the reflective image (B) but the man is not. Both are visible in the fused image (C). The fused image (C) puts the man into perspective relative to the fence.

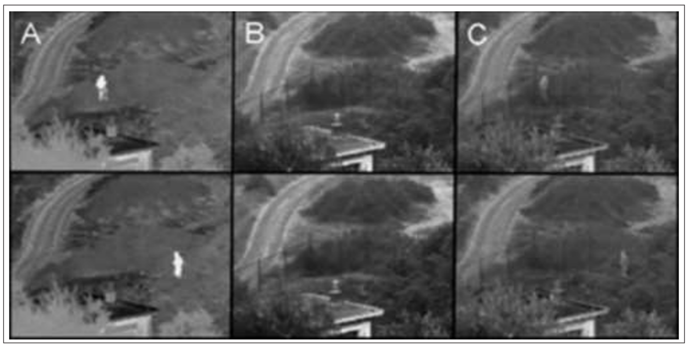

Figure 8: This Imagery is part of NATO_Camp_Sequence, available at https://Figshare.com. The Fused Images show the relationship between the man and fence, which is not discernable in either the Thermal or Reflective Images

Figure 9A, B, and C show thermal, reflective, and fused images, respectively. The top row shows images taken at night after a sunny day of insolation. The scene is dark to the eye, but a distant streetlight provides plenty of light for the image intensifier. The image shows a person walking near miniature cows. The person and cows are highlighted in the thermal image but subdued in the reflective image; MF enhanced the grayscale of the entire scene. As shown in Figures 7 and 8, scene objects that are washed out in one spectral band and too dark in the other tend to exhibit improved quantization in the fused result.

In Figure 9D, the person has a flashlight and is examining the gate. The people, cows, and vegetation are all visible, but the lights are not. In E, the reflective image, the lights are visible, but the person and cows are dim, whereas in the fused image F, both the lights and person are visible, and the scene content has improved grayscale because of the fusion.

Figure 9G, H, and I demonstrate that pointing the I2 camera at the flashlight dims the person and cows in the reflective image, whereas the fused image restores the scene.

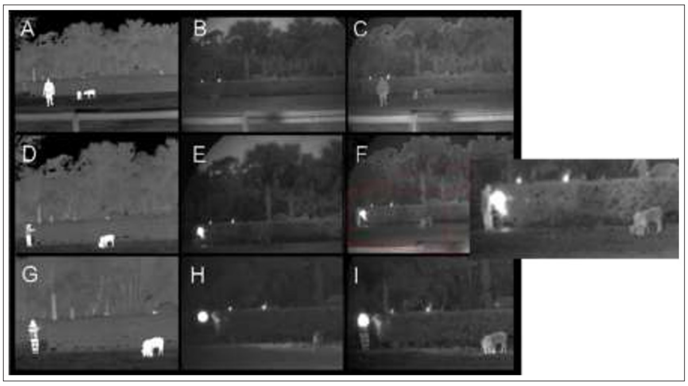

Figure 9: Images Captured on A Ranchette Featuring People and Miniature Cows

The images in Figure 10 were captured at night after a rainy day, and the thermal contrast is poor. However, the distant streetlight provides ample light for the image intensifier. Figure 10A, B, and C show thermal, reflective, and fused images, showing that the thermal imager automatically adjusted the contrast on the person. In the fused image shown in D, the thermal camera gain was reduced to lower the thermal noise. The resulting image shows the reflective image details and enhances the rendition of the person.

Figure 11 shows the fusion of thermal and reflective images but uses color to highlight the thermal image data. The reflective image was converted into color by making each of the three-color slices equal to the monochrome grayscale image, and the thermal image was then fused with the red slice.

The thermal image in Figure 11 used the camera shown in Figure 4 but was down sampled to 128 × 98 pixels and then interpolated to the image intensifier resolution before fusion. Only one high-resolution image was required for the MF to produce a high-resolution fused image. In this example, the thermal camera had a low resolution.

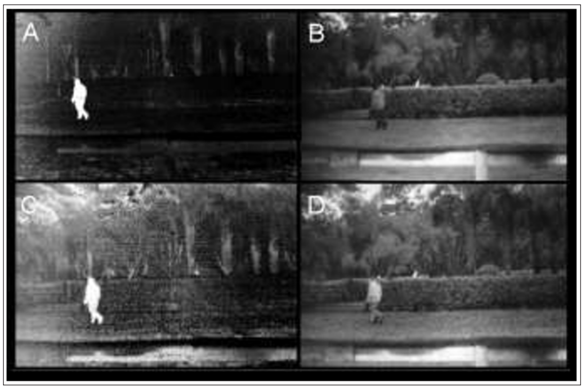

Figure 10: Contrast is Poor because these Images were taken at Night after a Rainy Day

Figure 11: The Thermal Image was down Sampled to 128 × 98 and then Interpolated Back to a High Resolution Before Fusion with the I^2 Image

Figure 12 shows a fused computed axial tomography (CAT) scan and magnetic resonance imaging (MRI) image. Note that the features of both images appear in the fused image, and the CAT scan grayscale improves.

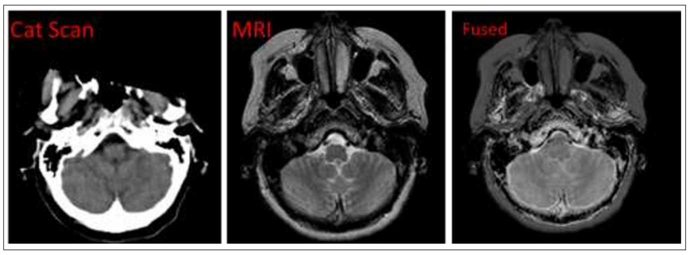

Figure 12: A CAT Scan, an MRI Image and a Fused Image. Images Taken from:
https://github.com/xianming-gu/Havard-Medical-Image-Fusion-Dataset

Figure 13A shows three thermal images, and B shows color map conversions based on the hue, saturation, and value (HSV). The problem with color-coding of intensity is that the structure is lost; therefore, intensity (or height, or whatever parameter is color-coded) is clarified by the coding, but the shape of the coded area is lost or at least compromised.

The images in Figure 13C are the fusions of those in A and B. The color mapping in B degrades image contrast and, in certain areas, makes the structure in the original picture difficult to see. Figure C illustrates that the fusion restores the image structure, making both the color map and object shape visible. Instead of the image grayscale equaling the heat intensity, fusion colors the spatial features based on the heat value.

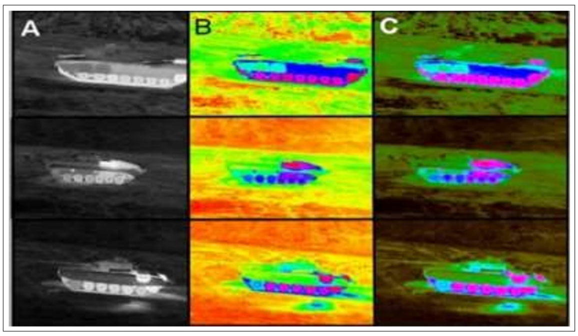

Figure 13: Fusing an HSV Color Map with the Original Restores the Structure of the Object

Conclusions

Currently, no experimentally validated objective metric exists for quantifying the effectiveness of fusion algorithms. However, based on our data collection, MF has proven to be effective both indoors and outdoors, in good and poor weather conditions, and in applications involving both tactical vehicles and animals with distinct thermal and reflective signatures.

These results suggest that the human visual system can extract salient features from both thermal and reflective images when they are fused via MF. MF enhances reflective scene comprehension by highlighting thermally significant elements such as humans, animals, and heated objects. In addition to combining thermal and reflective features, MF does not introduce artifacts or otherwise disrupt visual interpretation of the scene.

MF is computationally simple to implement and produces artifact-free imagery, meaning that features appear in the fused output only if they are present in at least one of the source images. Across all examples shown in Section 3, the thermal and reflective features are combined in a way that supports cognitive interpretation.

For instance, in Figure 5, the thermal signatures of the engine and exercised tracks add visual cues for vehicle identification; however, structural details are visible in the reflective image and not in the thermal image. In contrast, both the structure and heating are visible in the fused images. Similarly, in Figure 8, determining the person's position relative to the fence is only possible using the fused image. In Figure 6, the kneeling man is all but invisible in the image-intensifier image. In Figure 7, most objects exhibit low contrast in the thermal image, and the person behind the chair in the living room is washed out in the reflective image. The fused image brings out thermal grayscale details such as the woman and the metal stud signature on the living room wall behind the woman. In the ranchette images, fusion highlights the animals and humans, and lights are visible but do not hide objects behind glare. Figure 12 illustrates both the CAT scan and MRI features appearing in the fused image and the improved CAT scan feature grayscale.

If a color display is available, then the thermal and reflective images can be differentiated by color; in Figure 11, this makes the person and animals stand out. Note that a high-resolution thermal image is not needed to add thermal cues to the reflective picture. Before interpolating up for fusion, the thermal image only had 128 by 98 pixels.

MF might be useful in the medical imaging application. Only one example fusing CAT scan and MRI images was presented because the author does not have the skills needed to ascertain whether the fusion of medical images was effective.

As illustrated by Figure 13, MF also improves color map conversions by colorizing spatial features rather than mapping intensity to color. Fusion restores the tactical vehicle structure that is lost in the HSV conversion while maintaining the HSV color representation of intensity.

Color adds a capability to our vision beyond grayscale. However, colorizing an image by mapping intensity to color does not take advantage of color vision, because the spatial features conveyed by grayscale are lost. MF adds color features to grayscale, such that the monochrome features are not lost with the addition of color. Adding color features to grayscale is a second way that MF might have application in the medical imaging field.

MF is simple to implement, and the resulting fused imagery combines thermal and reflective visual cues to support the cognitive interpretation of the scene. An additional advantage is that the thermal camera need not be high resolution to add thermal cues to the reflective imagery. Also, MF might find application in the medical imaging field, either by fusing images gathered using different technologies, or by adding color to expand grayscale.

Funding

This study received no specific grants from any funding agency.

Conflict of Interest

The author declares no conflict of interest.

Data Availability Statement

Data available upon reasonable request.

References

1. Gouras Peter, Kolb H, Fernandez E, Jones B (2005) Webvision: The Organization of the Retina and Visual System. Color Vision University of Utah Health Sciences Center https://pubmed.ncbi.nlm.nih.gov/21413389/.
2. Gouras Peter (2025) The Organization of the Retina and Visual System. Color vision https://webvision.med.utah.edu.
3. Lee Barry B (2019) Sensitivity to chromatic and luminance contrast and its neuronal substrates, Current Opinion in Behavioral Sciences. 30: 156-162.
4. McCann, John J (2017) Retinex at 50: color theory and spatial algorithms, a review. Journal of Electronic Imaging 26: 31204.
5. Heidi Hofer, Joseph Carroll, Jay Neitz, Maureen Neitz, David R Williams (2005) Organization of the human trichromatic cone mosaic. JNEUROSCI 35-46.