May 02, 2019

# Improving Low-Light Imaging by Using Infrared and Software Interpolation

Kavinaath Murugan

Aaditya Kandibanda

# Improving Low-Light Imaging by Using Infrared and Software Interpolation

**Abstract**:

Complementary metal-oxide-semiconductor (CMOS) sensors are found in many imaging systems, such as cameras, and can be tuned to receive various ranges of wavelengths of light. In low-visible-light conditions, a camera with a CMOS sensor tuned to visible light, as is common in many portable electronic devices, may struggle to produce clear or detailed images. Likewise, even though a camera with an infrared-tuned CMOS sensor may detect additional information unavailable to the visible-light CMOS sensor in low-visible-light conditions, the resulting infrared-only image would lack an accurate color representation of the scene. Nevertheless, an imaging system featuring a machine-learning computer-aided vision system configured to blend color data from a visible-light-tuned CMOS sensor and morphological or topographical data from an infrared-tuned CMOS sensor can produce clear, colorful, and detailed images.

**Keywords:**

Camera, lens, image capturing device, digital single-lens reflex camera, image sensor, image detector, ambient light sensor, photodiode, visible spectrum, infrared, machine learning, neural network, deep learning, artificial intelligence, AI, composite-image.

**Background:**

Digital cameras are found in many different types of electronic devices and can be used to capture a variety of still images and motion videos. However functional and ubiquitous they may be, all digital cameras are limited by the amount of light or electromagnetic radiation they can receive and collect using a sensor. In low-light conditions, software components that manage the light collection for a digital camera struggle to collect the proper amount of light to produce a clear and detailed image. Simply collecting additional light over a longer period of time may produce a blurry final image, especially in handheld electronic devices. Outside of a controlled studio environment, the introduction of additional light, such as with a flash, may produce a final image with an unbalanced light or color profile because the flash provides a disproportionate amount of light to part, but not all, of the scene in the captured image. For some digital camera users, the strong highlights and unnatural result within portions of a final image taken with a flash may not be an improvement over a darkened image taken in low-light conditions without the flash.

**Description:**

Complementary metal-oxide-semiconductor (CMOS) sensors are found in many imaging systems, such as cameras, and can be tuned to receive various ranges of wavelengths of light. In low-visible-light conditions, a camera with a CMOS sensor tuned to visible light, as is common in many portable electronic devices, may struggle to produce clear or detailed images. Likewise, even though a camera with an infrared-tuned CMOS sensor may detect additional information unavailable to the visible-light CMOS sensor in low-visible-light conditions, the resulting infrared-only image would lack an accurate color representation of the scene. Nevertheless, an imaging system featuring a machine-learning computer-aided vision system configured to blend color data from a visible-light-tuned CMOS sensor and morphological or topographical data from an infrared-tuned CMOS sensor can produce clear, colorful, and detailed images.

Consider the dilemma faced by a photographer illustrated in Figure 1. Here, the photographer is attempting to capture the Taj Mahal in a low-light condition, such as early in the morning or late in the evening. Unless the photographer intends to produce a quasi-silhouette-style image, the final image produced by the photographer's visible light camera lacks sufficient detail to be interesting but too much detail to be a true silhouette. Given the size of the Taj Mahal, the addition of a flash will be unlikely to change the overall final image to any significant degree. Without some other way to capture more information, or interpret the available information captured, the photographer may be disappointed by the final result.

**Figure 1**

Instead, consider the imaging system of Figure 2 featuring a machine-learning computer-aided vision system (hereinafter referred to as the computer vision camera or vision system) configured to blend color data from a visible-light-tuned CMOS sensor and morphological or topographical data from an infrared-tuned CMOS sensor. This computer vision camera can significantly improve the final images obtained by the photographer. Many common elements of modern digital cameras or digital-camera equipped devices, such as memory, screens, processors, lenses, and the like, are omitted from Figure 2 for clarity and simplicity.

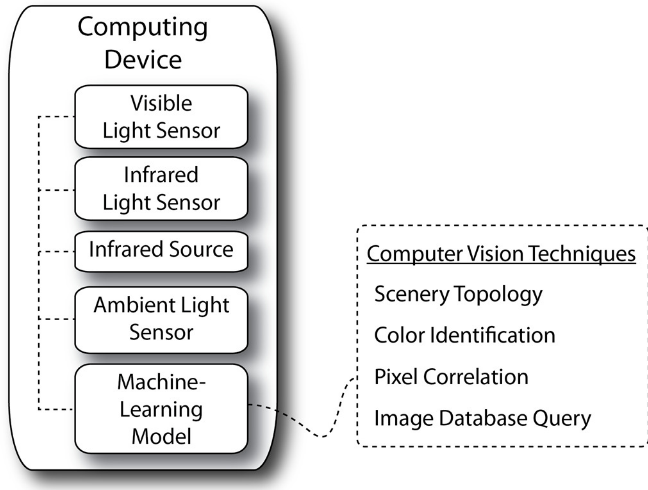


**Figure 2**

As shown in Figure 2, the machine-learning computer-aided vision system includes a visible light sensor (such as a CMOS sensor tuned to visible, colored light), an infrared light sensor (such as a CMOS sensor tuned to infrared light), an infrared source to supplement infrared light naturally occurring in a scene, an ambient light sensor, and a machine-learning model. In some cases, the ambient light sensor can be a standalone or a separate device, such as a light meter. In other cases, the visible light sensor can perform the light receiving and recording functionality of a CMOS sensor and the general light condition analysis of a light meter. The infrared sensor functions to receive infrared light or infrared electromagnetic radiation, which is almost always present in the physical world. The infrared source, such as a bulb or a strobe, can add additional infrared light to a scene to improve the performance of the vision system. The ambient light sensor can assist the vision system in determining when or how much infrared light to introduce into the scene with the infrared source.

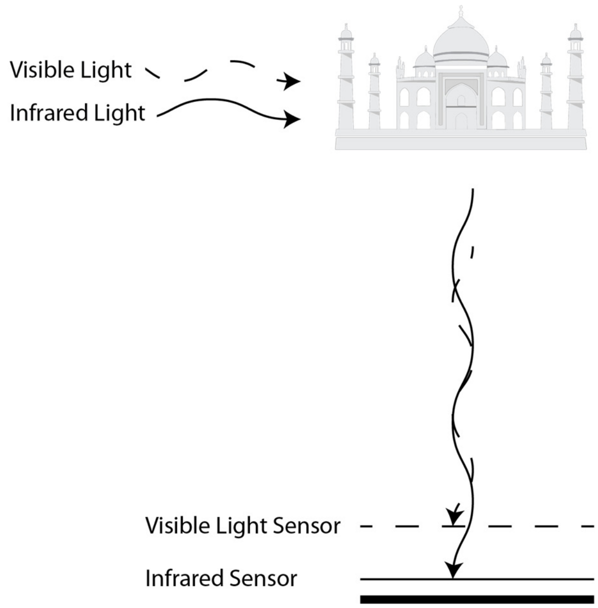


**Figure 3**

In some arrangements, the visible light sensor and the infrared sensor are stacked one on top of another or in series. Figure 3 illustrates the visible light sensor between the infrared sensor and the Taj Mahal scene being captured. Other arrangements could reverse the relative positions of the sensors depending on hardware or other configuration limitations. Although a stacked arrangement may be preferred because of reduced difficult or complexity in correlating information received at the multiple sensors, geometric and other physical limitations may not allow for this preferred arrangement.

As shown in Figure 4, the visible light sensor and the infrared sensor may be positioned adjacent one another to accommodate some devices equipped with the vision system. The machine-learning model may correlate light received at the visible light sensor with infrared light received at the infrared sensor. In some aspects, this arrangement and correlation of received light may be preferred because it allows for each sensor to have an unobstructed view of the scene.

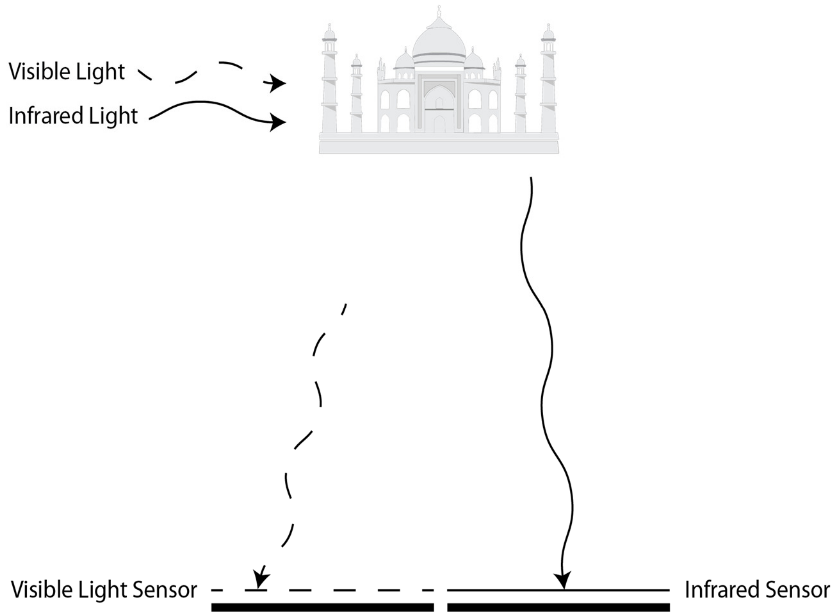


**Figure 4**

The vision system described herein can improve the final image obtained in the low-light conditions of Figure 1 by combining various sets of data that can be correlated and blended together to produce a final composite-image. Figure 5 illustrates an advantage of collecting infrared light in low-light conditions. Infrared light or electromagnetic radiation is often known as "heat radiation". Thermal cameras may acquire infrared light and represent this as the heat profile of a scene. Different types of materials or compositions of matter release, absorb, or reflect heat in different manners. For this reason, infrared light can be particularly useful in low-visible-light imaging, especially in the acquisition of edges where material composition changes from one material to another. Additionally, alterations or variations in geometry or shape may also affect the release, absorption, or reflection of heat or infrared radiation.

As shown in Figure 5, edges of portions of the Taj Mahal like the turrets or main façade, release or reflect heat differently than the air spaces in-between them. This difference can be read as an edge. Further, different substances in a continuous surface, such as different materials of construction like stone, glass, paint, or metal, or different physical shapes of a continuous surface, such as curves, sharp edges, areas in shadow, or areas exposed to light, can all release or reflect heat or infrared light in different manners. This variety of differences can be collected by the infrared sensor and reveal detail invisible to the visible light sensor or the human eye. Figure 5 illustrates how the opaque, silhouette-style Taj Mahal of Figure 1, can now be viewed in clear outlines and detail.

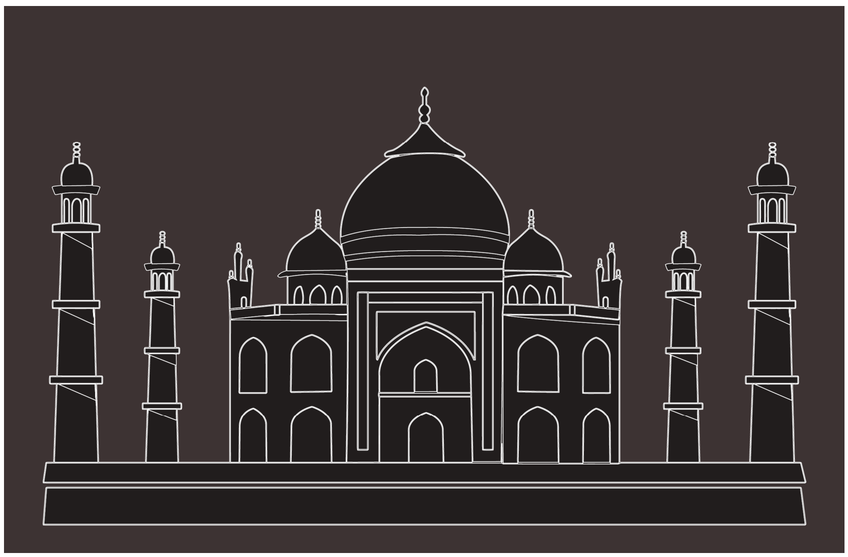

**Figure 5**

Although useful for providing additional detail, the infrared sensor alone does not produce a desirable final image. Layering the image of Figure 5 onto the image of Figure 1 would not appreciably improve the final image nor result in something substantially different from Figure 5. But, this vision system includes a machine-learning model that can correlate and interpolate additional information. Although the low-light conditions of Figure 1 do not produce enough light to capture a pleasing final image with the visible light sensor alone, the image of Figure 1 includes gradients, variations, and differences among the received visible light. The image also includes variations in color that, to a computing system, are different enough to separate and categorize even if the human eye may not differentiate among the colors based on these subtle differences.

The vision system identifies regions of a similar infrared signature with regions in the image acquired by the visible light sensor. The system interprets a similar infrared signature as sharing a similar color or composition. For areas of the image obtained by the visible light sensor for which insufficient or less color data was collected that share a similar infrared profile, the vision system interpolates and color these areas, based on those portions of the image obtained by the visible light sensor for which sufficient or more color data was collected.

Figure 6 illustrates different shades of color only slightly visible in Figure 1 that are clarified and layered on the edges detected in Figure 5. As seen in Figure 6, the vision system supplements colors by identifying a similar type of infrared signature and determining that similar signatures share similar colors even if the visible light sensor did not pick up the color well-enough in the location because of the low-light conditions.

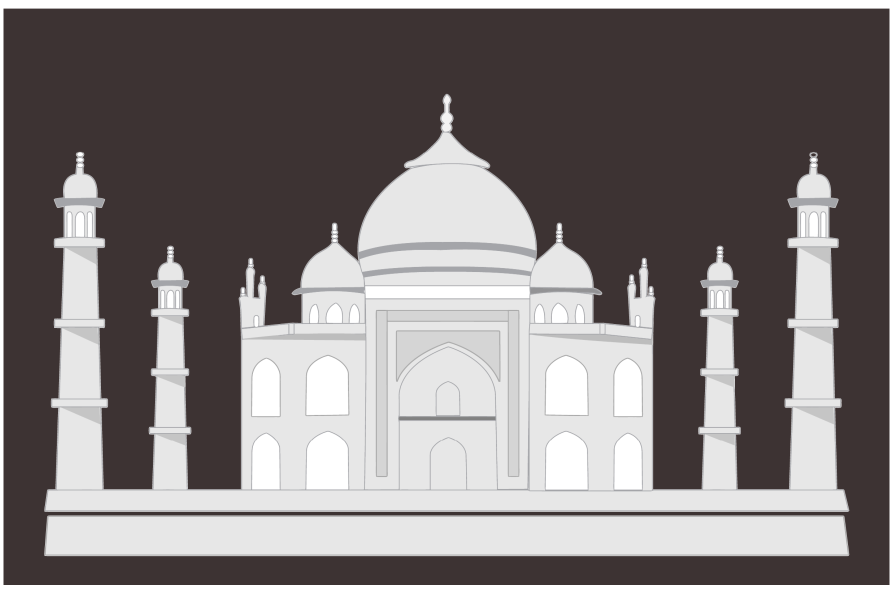

**Figure 6**

Figure 6 also illustrates an additional feature of the vision system. Sometimes, the combination of infrared and visible light may not be enough to produce a desirable final composite-image. The vision system analyzes the acquired information, makes a determination about the contents of the low-light acquired image, and queries an image database for other images of similar subject matter. For example, there are many images of the Taj Mahal acquired in better-light conditions. The vision system may supplement the data acquired by the infrared and visible light sensors with additional color or edge information from existing images of the Taj Mahal.

In other circumstances, the vision system may query other images of individuals (e.g., from an existing album on the computing device) to supplement images of the individuals acquired in low-light conditions. For example, a couple on a tour of the Taj Mahal may have taken a picture at a visitor's center, presumably with better lighting conditions, prior to requesting a picture of themselves with the Taj Mahal in the background. Because the images are taken within a relatively short or related time period, the system may interpret the image of the couple at the visitor's center as a good candidate for information to supplement an image taken in lower-light conditions soon thereafter. The described vision system described can query one set of images to supplement the Taj Mahal and a separate set of images to supplement the individuals in the construction of a final composite-image.

In low-visible-light conditions, a computing device equipped with a sensor tuned to visible light and a sensor tuned to infrared light can acquire substantial information about a scene. A machine-learning computer-aided vision system analyzes the captured scene, supplements and refines the acquired visible and infrared information by acquiring supplement information from a known-image database, and produces a desirable final image by blending the captured and acquired information into a final composite-image.