Chapter 1
Motivation

Currently, humans live in a digitally connected community surrounded by exploding information from numerous multimedia sources. It seems as if everything is available through the internet. For example, search for a paper using Google, buy a book from Amazon, or download images and video collected from digital cameras. Even for the health community, there are many choices of imaging solutions, such as x-rays, ultrasound, computed tomography (CT), and magnetic resonance imaging (MRI). For biometrics, fingerprint-, iris-, and face-recognition techniques are available. To execute surveillance, there are multispectral (including thermal) imaging systems for nighttime and electro-optical visual imaging for daytime. There appears to be significant data access to any form of imagery, so what is the need for image fusion and how does it enhance imagery? This chapter answers these questions with a look at four applications: face recognition, medical diagnosis, night-vision scene enhancement, and multifocus rendering.

1.1 What Is Image Fusion?

Multisensor image fusion is the process of combining relevant information from two or more images into a single image.[1] The prevailing thought is that the resulting image will be more informative than any of the input images.[2]

The multiple input images can be multimodal, multifocus, or multiview images. For example, CT and MRI images, and visible (RGB) and thermal (e.g., long-wave infrared) images, are obtained using different imaging modalities, which constitute *multimodal* images. Multifocus images are captured by focusing on objects at different distances. Stereo images taken at different viewing angles form multiview images. Part 1 of the book develops image fusion concepts (Chapters 2–4).

Image fusion is the process of combining multisource images. Averaging two input images pixel by pixel is a classic, simple fusion process, although it may smooth the image features between images. The image fusion process can be enhanced by using a set of image transforms and fusion rules, which will be elaborated in Part 2 on image fusion theory (Chapters 5–8).

The resulting image from a fusion process is called a *fused image*. The fused image should contain more comprehensive information than the input images alone. Objective metrics and subjective evaluations are used to measure and evaluate the fused image, as presented in Part 3 on image fusion evaluation (Chapters 9 and 10).

Image colorization is a special application of image fusion. For example, multispectral night vision (NV) images are first fused and then painted with natural colors. The colorized NV images look like daylight pictures, which human operators can utilize for decision making. Part 4 examines image fusion applications (Chapters 11 and 12).

Chapter 13 provides a summary of the book details.

1.2 The Purpose of Image Fusion

Information fusion exists naturally in biological vision systems. For example, many animals have two eyes for binocular stereo vision to aid in capturing prey. Of the five human senses, the two eyes and two ears are used for multimodal and multichannel information fusion. Fusion mechanisms also exist in other biological vision systems, such as that of snakes, which have retinotopically mapped infrared sensors on their visual sensors. Image fusion is a special case of information fusion, where the source information includes different images.

There are many advantages of using image fusion:

1. *Multimodal* image fusion provides comprehensive information that is useful for pattern recognition and visual inspection. For example, a combined CT and MRI image presents complete information about both bone structure and soft tissues; and multispectral image fusion can greatly benefit target detection, tracking, and recognition.
2. *Multichannel* image fusion can increase the signal-to-noise ratio (SNR). Given a pair of stereo images (e.g., from a 3D camera), an average of two images will create a clearer image. Compare the reading of a book with two eyes versus one eye: the former is more comfortable (e.g., causes less fatigue) because humans can fuse the image signals from two eyes.
3. *Multiview* image fusion can increase the decision confidence, such as when recognizing a 3D object. A 3D model may be created from a pair of stereo images or from several multiview images (e.g., front, top, and side view). Regarding a face recognition application, suppose that three poses of facial images are collected onsite, e.g., front, left-half-profile, and right-half-profile images. If the recognition of three poses of facial images confirm the same person, then the likelihood and confidence of making a correct conclusion is much higher.

Typical applications of fused images can improve the performance of pattern recognition. For example, multispectral image fusion can help improve the accuracy of target, object, or scene recognition.

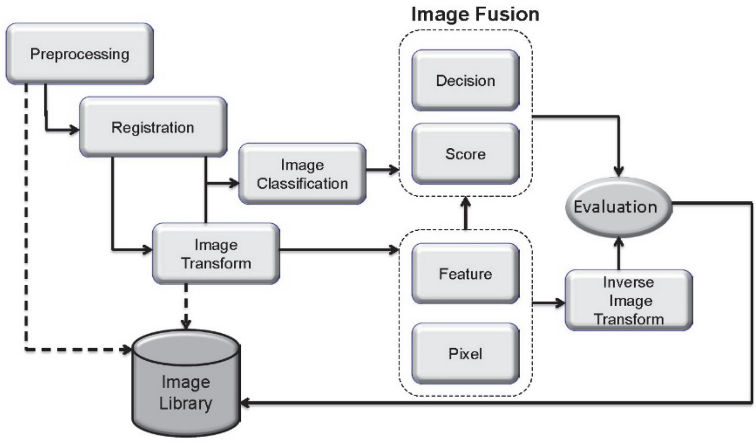


Figure 1.1 Typical image fusion process.

1.3 How Image Fusion Works

A typical image fusion process consists of the following steps (see Fig. 1.1):

1. *preprocessing*, which includes denoising, normalization, and enhancement;
2. *registration*, which aligns source images and is required by image fusion at the pixel level;
3. *transform*, which coverts images to coefficients, such as a wavelet transform;
4. *fusion*, which combines pixels or coefficients according to set fusion rules;
5. *invert transform*, which coverts the coefficients to images; and
6. *evaluation*, which evaluates the quality of the fused image.

Given a pair of registered CT and MRI images, for instance, a simple fusion process can be implemented by averaging two images pixel by pixel. Averaging is a fusion rule because there is there no transform (or inverse transform) applied to the images. A weighted average or principal component analysis (PCA) can be used instead of averaging to adaptively improve the results. The fused images from the CT and MRI may be evaluated by radiologists, augmented with objective metrics (e.g., entropy), to assist decision making (e.g., cancer detection). This book covers all fundamental aspects of the image fusion process, with an emphasis on image fusion methods and evaluation.

1.4 Applications

Several image fusion examples are briefly mentioned here to illustrate image fusion advantages.

(a) (b) (c) (d)

Figure 1.2 Fusion of facial images: (a) Visible image, (b) thermal image, (c) image fused by a Laplacian pyramid, and (d) orientation-based fusion.

1.4.1 Face recognition

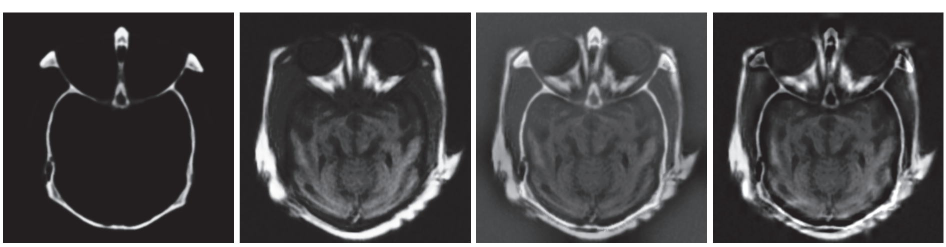
Thermal face recognition is useful for poor illumination or at nighttime. However, long-wave infrared (LWIR) light cannot pass through glasses (Fig. 1.2(b)). The fusion of two spectral images provides the mixed features shown Figs. 1.2(c–d), where the eyes are visible. The experiments proved that the electro-optical (EO) and infrared (IR) fused facial images could be used to improve the recognition accuracy.[3]

1.4.2 Biomedical applications

As shown in Figs. 1.3(c–d), the fused images can present more complete information by combining CT and MRI images. The fused images are useful for surgery planning, tumor detection and location, and biological change detection.

1.4.3 Visual inspection

Imagine a scenario where surveillance is conducted at night with an unmanned aerial vehicle (UAV) equipped with multispectral imaging devices, as shown in Figs. 1.4(a–b). The UAV pilot may have to watch two channel images separately and synthesize the information manually. The fused image in grayscale (Fig. 1.4(c)) is useful for image segmentation and scene classification. The colorized image shown in Fig. 1.4(d) looks like a daylight picture, which is more easily and comfortably interpreted by human operators.



(a) (b) (c) (d)

Figure 1.3 Fusion of medical images: (a) CT, (b) MRI, (c) image fused by a Laplacian pyramid, and (d) image fused by the advanced DWT.

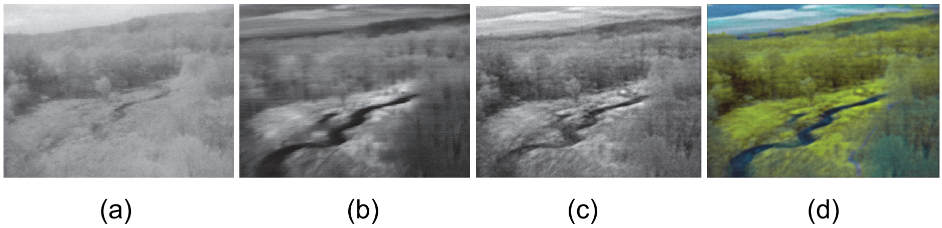


Figure 1.4 Segmentation-based NV colorization: (a) and (b) NV images (intensified and thermal), (c) image (grayscale) fused by the advanced DWT, and (d) the colorized image.

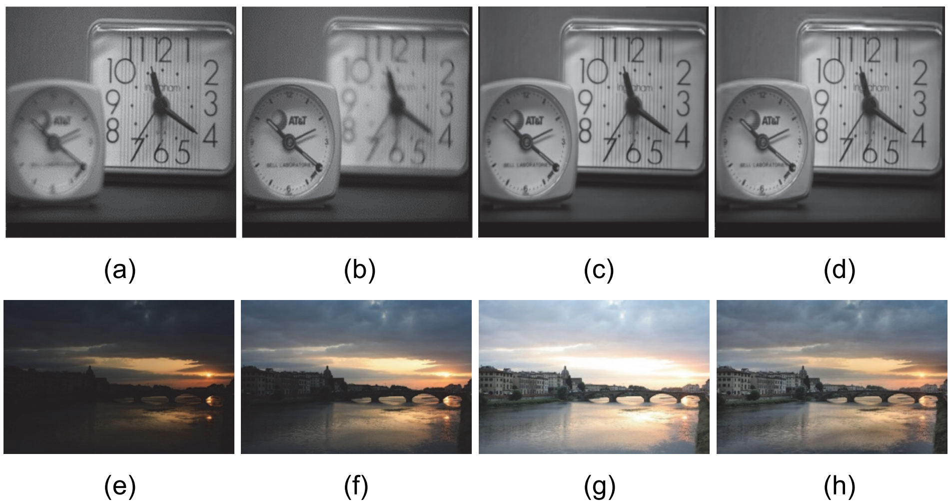


Figure 1.5 Image fusion examples: (a–b) Multifocus images, (c) image fused by a contrast pyramid, (d) image fused by the advanced DWT, (e–g) three exposure images (low, medium, high), and (g) image fused by Paul et al.'s method.[4]

1.4.4 Multifocus and multiexposure fusion

Multifocus images are captured using different focus points due to the limited depths of field (DOFs) of cameras. Conversely, multiexposure images are captured by setting different exposure degrees due to the high dynamic range in a scene. As shown in Fig. 1.5, image fusion techniques can produce all-in-focus images and all-well-exposed images.

1.5 Summary

Image fusion combines multisource images into one image. The fused image is derived to provide more comprehensive information, such as the performance improvement of scene characterization, pattern recognition, or human action (e.g., target detection and tracking). With more advanced imaging modalities, pragmatic image fusion necessitates an understanding of the techniques, evaluation measures, and appropriate applications. However, the development of an image fusion process for a particular application requires deep understanding of image fusion concepts, methods, and evaluations. This book presents perspectives on how to conduct, use, and measure image fusion in four parts.

The rest of the book is organized into four sections based on multimodal, multiview, and multifocus examples. The sections include a discussion of the concepts, theory, evaluation, and development, as shown in Fig. 1.6. Supporting datasets, software (code), and research papers are made available on the authors' websites for further use by the image fusion community. Table 1.1 summarizes the methods highlighted in this book, along with their advantages and disadvantages.

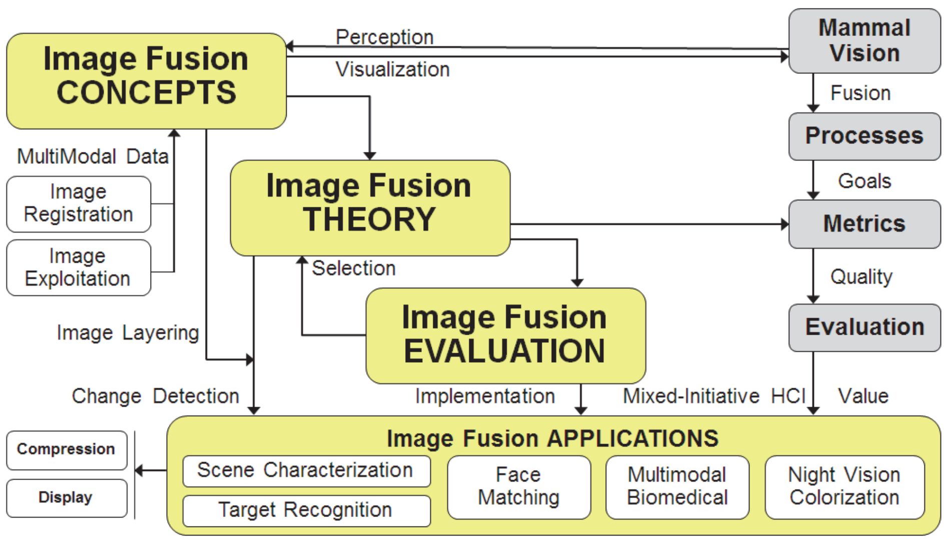


Figure 1.6 Book overview.

Table 1.1 Overview of fusion methods.

Chapter	Method	Advantage	Limitation	Application
2	Feature-level fusion	Affords association in multiresolution/ multimodal images	Does not account for denoising	Target detection with EO/IR
3	Opponent processing	Combines local features similar to biological systems	Focused on object detection and not whole-image analysis	Target detection in EO/IR for night vision
4	Bayes network fusion	Captures environmental conditions such as lighting	Requires a prior modeling of at-image capturing	Dynamic tracking in EO/IR sequential images
5	Decision fusion	Combines several decisions from different sources	Does not weigh input decisions	Face recognition
6	Score fusion	Affords weighting of multimodal scores	Selection of scoring logic can vary depending on image types	Face recognition
7	Wavelet fusion	Supports multifocus/ multiresolution fusion	Whole-image fusion, where selection can be qualitative	Face recognition
8	Histogram matching	Objective-based methods are repeatable	Analysis can vary between human observers	Colorization
9	Objective evaluation index	Comprehensive metric incorporating phase, gradient, and contrast	Metrics have not been confirmed to correlate with human observers	Colorization
10	Relative rate increase	Repeatable and applicable to dynamic systems	Does not capture qualitative aspects	Colorization, Night vision, medical
11	Stereo fusion	Captures multimodal and multispectral analysis	Focused on 2D, but can be extended to 3D	Face recognition
12	Iterative fusion	Assists in analysis refinement to support user needs	Requires stopping and confirmation from humans	Medical, concealed-weapon detection

References

1. M. B. A. Haghighat, A. Aghagolzadeh, and H. Seyedarabi, "Multi-focus image fusion for visual sensor networks in DCT domain," *Comp. Elect. Engineer.* **37**(5), 789–797 (2011).
2. M. B. A. Haghighat, A. Aghagolzadeh, and H. Seyedarabi, "A non-reference image fusion metric based on mutual information of image features," *Comp. Elect. Engineer*. **37**(5), 744–756 (2011)
3. Y. Zheng and E. Blasch, "The advantages of stereo vision in a face recognition system," *Proc. SPIE* **9091**, 90910Y (2014) [doi: 10.1117/12.2049985].
4. S. Paul, I. S. Sevcenco, and P. Agathoklis, "Multi-exposure and Multi-focus image fusion in gradient domain," *J. Circuits, Systems, Computers* **25**(10) (2016).