THEORY AND APPLICATIONS OF IMAGE REGISTRATION

ARTHUR ARDESHIR GOSHTASBY

with website

WILEY

Theory and Applications of Image Registration

Theory and Applications of Image Registration

Arthur Ardeshir Goshtasby

WILEY

This edition first published 2017

Registered Offices
John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA

Editorial Office
111 River Street, Hoboken, NJ 07030, USA

For details of our global editorial offices, customer services, and more information about Wiley products visit us at www.wiley.com.

Wiley also publishes its books in a variety of electronic formats and by print-on-demand. Some content that appears in standard print versions of this book may not be available in other formats.

Limit of Liability/Disclaimer of Warranty
The publisher and the authors make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties; including without limitation any implied warranties of fitness for a particular purpose. This work is sold with the understanding that the publisher is not engaged in rendering professional services. The advice and strategies contained herein may not be suitable for every situation. In view of on-going research, equipment modifications, changes in governmental regulations, and the constant flow of information relating to the use of experimental reagents, equipment, and devices, the reader is urged to review and evaluate the information provided in the package insert or instructions for each chemical, piece of equipment, reagent, or device for, among other things, any changes in the instructions or indication of usage and for added warnings and precautions. The fact that an organization or website is referred to in this work as a citation and/or potential source of further information does not mean that the author or the publisher endorses the information the organization or website may provide or recommendations it may make. Further, readers should be aware that websites listed in this work may have changed or disappeared between when this works was written and when it is read. No warranty may be created or extended by any promotional statements for this work. Neither the publisher nor the author shall be liable for any damages arising here from.

Library of Congress Cataloging-in-Publication Data

Names: Goshtasby, Ardeshir, author.
Title: Theory and applications of image registration / Arthur Ardeshir Goshtasby.
Description: Hoboken, NJ : John Wiley & Sons, 2017. | Includes bibliographical references and index. | Description based on print version record and CIP data provided by publisher; resource not viewed.
Identifiers: LCCN 2017007704 (print) | LCCN 2017012539 (ebook) | ISBN 9781119171720 (PDF) | ISBN 9781119171737 (ePub) | ISBN 9781119171713 (cloth)
Subjects: LCSH: Image registration.
Classification: LCC TA1632 (ebook) | LCC TA1632 .T4815 2017 (print) | DDC 006.6–dc23
LC record available at https://lccn.loc.gov/2017007704

Cover image: © Hayami Yanagisawa / EyeEm/Gettyimages
Cover design by Wiley

Set in 10/12pt WarnockPro by SPi Global, Chennai, India

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

To my wife Mariko, and my parents Bahram and Firoozeh.

Brief Table of Content

Contents

Contributors

EDGARDO MOLINA received his PhD in computer science from the CUNY Graduate Centre in New York in 2015 and a BS degree in computer science from the Macaulay Honors College at The City College of New York – CUNY in 2005. His research interests are computer vision, human–computer interaction, 3-D visualization, and applications for assistive technologies. He has been cofounder and CEO of Vista Wearable, Inc., developing wearable assistive devices for the visually impaired and providing consulting services to other corporations.

WAI L. KHOO received his PhD in computer science from the Graduate Centre – CUNY in 2016 and BS and MS degrees in computer science at The City College of New York – CUNY in 2009 and 2012, respectively. His research interests include computer vision, multimodal sensing, human–computer interaction, and virtual/augmented reality and also application of computer vision to assistive technology for the visually impaired.

HAO TANG is an assistant professor of computer science at Borough of Manhattan Community College – CUNY. He received a PhD degree in computer science in 2013 from the Graduate Centre – CUNY. He was an adjunct faculty member with The City College of New York and a senior researcher at The City College Visual Computing Research Laboratory. His research interests include 3-D computer modeling, visualization of large-scale scenes, image understanding, and mobile vision navigation, as well as their applications in surveillance, and assistive technology for the visually impaired. His research paper in the International Conference on Multimedia and Expo won the best paper award Finalist.

ZHIGANG ZHU is the Herbert G. Kayser professor of computer science at The City College of New York and the Graduate Center – CUNY. He is also the director of The City College Visual Computing Laboratory (CCVCL). His research interests include 3-D computer vision, multimodal sensing, virtual/augmented reality, video representation, and various applications of assistive technology in environment, robotics, surveillance, and transportation. He has published over 150 technical papers in related fields. He is an associate editor of the Machine Vision Applications Journal and has been technical editor of the ASME/IEEE Transactions on Mechatronics.

Acknowledgments

The satellite images used in this monograph are courtesy of NASA, USGS, Google, and Microsoft Bing; the aerial images are courtesy of Image Registration and Fusion Systems and the Computer Vision Group at Brown University; and the medical images are courtesy of Kettering Medical Center. Results reported in this monograph are from research conducted with funding from NSF, NIH, AFRL, Kettering Medical Center, and the Ohio Board of Regents. The contributions and support of these organizations toward preparation of this monograph are greatly appreciated. The author would like to recognize and thank Edgardo Molina, Wai Lun Khoo, Hao Tang, and Zhigang Zhu, who contributed a key chapter to this monograph. Finally, the author would like to thank Libby Martin for editorial assistance in preparation of this monograph.

A.A.G

About the Companion Website

This book is accompanied by a companion website:

http://www.imgfsr.com/WileyRegistrationBook.html

[If the website cannot be accessed for some reason, contact the author directly at arthur.goshtasby@gmail.com for the online materials relating to the book.]

1

Introduction

The field of image registration came to being even before the advent of digital media. Dressler, in a US patent in 1956 [1], disclosed an electronic image comparison device that could register images recorded on films. By providing the mechanism to translate and rotate one film over the other and by projecting the images to a subtraction device using a half mirror and a full mirror, he provided the means to register and compare images. Circuitry was provided to produce an output at any given time with its amplitude proportional to the degree of match between the images. While observing the amplitude of the output, the user was enabled to interactively register the images.

The first work in digital image registration goes as far back as 1963 in the Ph.D. dissertation of Lawrence Roberts at MIT [2]. By aligning a 2-D projection of a 3-D polyhedral scene to edges of known polyhedral objects, he was able to identify polyhedral objects in the scene via image registration.

Image registration as we know it today came to being by the groundbreaking work of Anuta [3]. Anuta developed an automated method for spatially aligning remote sensing images by the fast Fourier transform.

Image registration is the process of spatially aligning two images of a scene. The alignment process determines the correspondence between points in the images. If the images are taken by the same sensor but at different times, the correspondence process makes it possible to detect changes in the scene occurring between the times the images are obtained [4]. If the images are obtained by different sensors, the correspondence process makes it possible to combine information in the images, creating a richer source of information about the underlying scene [5, 6].

By registering frames in a video captured by a moving camera, it becomes possible to distinguish image differences caused by camera motion from image differences caused by object motion, making it possible to track moving

Theory and Applications of Image Registration, First Edition. Arthur Ardeshir Goshtasby.

objects in the video [7]. If the images represent different views of a scene, the correspondence process makes it possible to determine the geometry of the scene [8].

This monograph covers the fundamentals of digital image registration. Only feature-based image registration methods are covered. Although a variety of optimization-based methods have been developed that use image intensities to register images, such methods are often very time consuming and are influenced by image noise. Feature-based methods are more resistant to noise and are computationally very efficient.

Image registration methods that use feature points and feature lines will be covered. Feature points identify locally unique neighborhoods in an image, facilitating the correspondence process. In addition to its location, a feature point may have various feature values, characterizing the neighborhood of the point. Feature points have been referred to as *control points, point landmarks, key points, corners, centroids*, and *tie points* in the literature.

A feature line at the minimum has a position and an orientation. The position of a line is the point on the line closest to the origin. The orientation of a line is the angle the normal to the line makes with the x-axis. These are the parameters used to define a line in polar form. In addition to the position and orientation of a line, the midpoint, length, endpoints, intensities, and gradients along the line can be used as features to describe the line.

The process of spatially aligning two images involves resampling one image to the geometry of the other. The image that is kept unchanged is called the *reference image* and the image that is resampled to the geometry of the reference image is called the *test image*. Reference image has also been called *source image* in the literature, and test image has also been called *target image* and *sensed image* in the literature.

An example of image registration is given in Fig. 1.1. Reference image (a) is a Landsat Multispectral Scanner (MSS) image of Kalkaska County, Michigan,

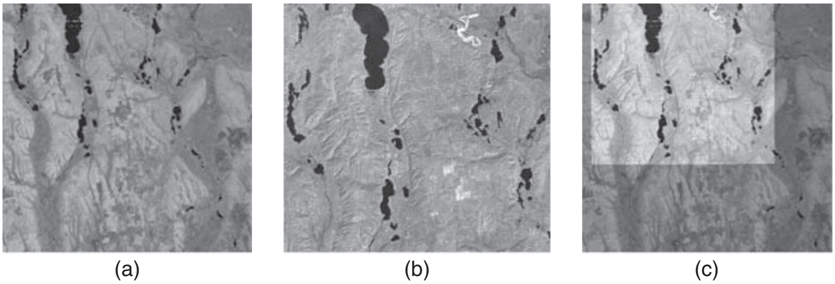


Figure 1.1 (a) A Landsat MSS image and (b) a Landsat TM image of Kalkaska County, Michigan. (c) The TM image is geometrically transformed to spatially align the MSS image. These images are courtesy of the USGS.

and test image (b) is a Landsat Thematic Mapper (TM) image of the same area. These images are courtesy of the US Geological Survey (USGS). The registered image (c) is obtained by geometrically transforming the TM image to spatially align the MSS image. By registering the images, it is possible to fuse information in the images or quantify differences between the images.

In this monograph, coordinates of points in the reference image will be denoted by (x, y), while coordinates of the corresponding points in the test image will be denoted by (X, Y). Image registration involves finding the relation between (x, y) and (X, Y). This relation is called a *transformation model.* A transformation model for registering 2-D images has two components, denoted by f_x and f_y:

$$X = f_x(x, y), \tag{1.1}$$

$$Y = f_y(x, y). \tag{1.2}$$

Functions f_x and f_y represent the x- and the y-components of the transformation. For each point (x, y) in the reference image, f_x and f_y determine the coordinates of the corresponding point (X, Y) in the test image. Therefore, once the components of a transformation model are determined, by scanning the reference image, for each pixel (x, y) there, the location (X, Y) of the same pixel in the test image is determined. Then, the intensity at (X, Y) in the test image is read and saved at (x, y) in a new image called the *resampled image.* The resampled image, therefore, has the geometry of the reference image and the intensities of the test image.

Note that functions f_x and f_y map pixels in the reference image to the corresponding pixels in the test image. This mapping in effect finds intensities in the test image corresponding to the pixels in the reference image. For integer pixel coordinates (x, y) in the reference image, f_x and f_y produce floating-point coordinates (X, Y) in the test image. The integers closest to X and Y are taken as the coordinates of the pixel in the test image corresponding to pixel (x, y). Alternatively, the intensity at location (X, Y) in the test image is estimated from the intensities of pixels surrounding (X, Y) by bilinear interpolation or cubic convolution [9].

1.1 Organization of the Book

This monograph provides theoretical basis and implementation details of methods for finding feature points/feature lines in images, establishing correspondence between the feature points/feature lines, and using the correspondences to compute the parameters of a transformation model to register the images.

Chapter 2 describes a method for determining the orientation of an image. An image containing a nonsymmetric pattern has a preferred orientation.

By determining the orientations of two images and by bringing the images to the same orientation, steps in image registration are simplified.

Chapter 3 discusses methods for identifying locally unique neighborhoods in an image. Centers of locally unique neighborhoods are taken as feature points. Rotation-variant and rotation-invariant methods are described. Rotation-variant methods are suitable for registering images that have the same orientation, and rotation-invariant methods are suitable for registering images that have different orientations.

Chapter 4 covers various line detection methods. Hough transform–based methods find lines among unorganized points, edge-based and curve-based methods find lines among edge contours, and region subdivision-based methods find lines among raw image intensities. Region subdivision-based methods by subdividing an image into regions of similar gradient directions and by fitting a line to each region detect lines.

Chapter 5 discusses the point correspondence problem. Given two sets of points, the problem of finding correspondence between points in the two sets is addressed. Corresponding points are also referred to as *homologous points*. First, clustering methods that use information about the locations of the points to find homologous points are described. Then, methods that use various information in addition to point locations to find homologous points are described. Finally, methods that determine homologous points in images by template matching are described. Geometric constraints and robust estimators that remove outliers from obtained correspondences are also discussed.

The problem of line correspondence is covered in *Chapter 6*. First, formulations that relate the parameters of homologous lines in two images to the parameters of a transformation model registering the images are derived. Then, methods that determine homologous lines in images by line grouping are discussed.

Chapter 7 covers topics relating to nonrigid image registration. Given a set of homologous points in two images of a scene, the problem of finding the parameters of an elastic transformation model to transform the geometry of the test image to the geometry of the reference image is discussed. Various elastic transformation models suitable for nonrigid image registration are described.

While Chapters 2–7 cover methods for the registration of projective images, *Chapter 8* covers methods for the registration of tomographic images. A tomographic image sequence represents a stack of cross-sectional images captured of a 3-D object. In this chapter, after discussing a method for converting a tomographic image sequence into an isotropic volume, methods for rigid and nonrigid registration of isotropic volumes are discussed.

In *Chapter 9*, methods for validating the accuracy of an image registration software are discussed. Specifically, methods that use a gold standard or simulation data to determine the accuracy of an image registration software are discussed. A number of methods that measure registration accuracy in the absence of simulation data or a gold standard are also discussed.

In *Chapter 10*, registration of video images to create panoramas, including stereo panoramas, is discussed. Registration under different camera motions, including pure rotation, pure translation, and general motion, are discussed. The structure of a camera setup to create panoramas in real time with a minimal computational requirement is also described.

Chapter 11 discusses registration of multitemporal images. It is found that feature lines remain more stable than feature points in images of scenes going through natural disasters. Lines representing roads, bridges, and edges of building are hardly affected by storms, flooding, and earthquakes. Fusion of registered images to detect scene changes in multitemporal images is also discussed.

In *Chapter 12*, open problems and research topics in image registration are mentioned. These include rotation-invariant similarity and distance measures for registration of multimodality images, rotation-invariant descriptors for registration of multimodality images, and the relation between parameters of a nonrigid transformation model and parameters of homologous lines in images.

The majority of algorithms discussed in this monograph have been implemented and the software packages producing the results reported in this monograph are made available to the readers as learning tools. The software packages may be downloaded from http://www.imgfsr.com/WileyRegistrationBook.html. *Appendix A* is meant to serve as a user's guide to the software packages.

1.2 Further Reading

Image registration has been an active area of research in image processing and computer vision since early 1970s. Papers on image registration regularly appear in *IEEE Transactions on Pattern Analysis and Machine Intelligence, IEEE Transactions on Image Processing, IEEE Transactions on Geoscience and Remote Sensing, IEEE Transactions on Medical Imaging, Medical Image Analysis, International Journal of Computer Vision, Image and Vision Computing, Pattern Recognition, Pattern Recognition Letters,* and *Computer Vision and Image Understanding* journals.

Surveys of image registration methods [10–18], special issues [19–22], and books on the subject [9, 23–28] also appear in the literature.

References

1 R. Dressler, *Image Matching Apparatus*, US Patent 2,989,890, Filed No. 13, 1956, Patented June 27, 1961.

2 L. G. Roberts, *Machine Perception of Three-Dimensional Solids*, Department of Electrical Engineering, Massachusetts Institute of Technology, 1963.

3 P. E. Anuta, Spatial registration of multispectral and multitemporal digital imagery using fast Fourier transform techniques, *IEEE Transactions on Geoscience Electronics*, **8**(4):353–368, 1970.

4 C. Zhao and A. Goshtasby, Registration of multitemporal aerial optical images using line features, *ISPRS Journal of Photogrammetry and Remote Sensing*, **117**:149–160, 2016.

5 A. Goshtasby and S. Nikolov, Image fusion: advances in the state of the art, *Information Fusion*, **8**:114–118, 2007.

6 F. Maes, D. Loeckx, D. Vandermeulen, and P. Suetens, Image registration using mutual information, in *Handbook of Biomedical Imaging*, Springer, 295–308, 2015.

7 M. Linger and A. Goshtasby, Aerial image registration for tracking, *IEEE Transactions on Geoscience and Remote Sensing*, **53**(4):2137–2145, 2015.

8 O. C. Ozcanli, Y. Dong, J. L. Mundy, H. Webb, R. Hammoud, and V. Tom, A comparison of stereo and multiview 3-D reconstruction using cross-sensor satellite imagery, in *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 17–25, 2015.

9 A. Goshtasby, *2-D and 3-D Image Registration for Medical, Remote Sensing, and Industrial Applications*, Wiley Press, 2005.

10 W. Crum, T. Hartkens, and D. Hill, Non-rigid image registration: theory and practice, *British Journal of Radiology*, **77**(2):140–153, 2004.

11 M. Holden, A review of geometric transformations for nonrigid body registration, *IEEE Transactions on Medical Imaging*, **27**(1):111–128, 2008.

12 J. A. Schnabel, M. P. Heinrich, B. W. Papież, and Sir J. M. Brady, Advances and challenges in deformable image registration: from image fusion to complex motion modelling, *Medical Image Analysis*, **33**:145–148, 2016.

13 A. Sotiras, C. Davatzikos, and N. Paragios, Deformable medical image registration: a survey, *IEEE Transactions on Medical Imaging*, **32**(7):1153–1190, 2013.

14 P. A. van den Elsen, E. J. D. Pol, and M. A. Viergever, Medical image matching: a review with classification, *IEEE Engineering in Medicine and Biology*, **3**:26–39, 1993.

15 M. A. Viergever, J. A. Maintz, S. Klein, K. Murphy, M. Staring, and J. P. W. Pluim, A survey of medical image registration, *Medical Image Analysis*, **33**:140–144, 2016.

16 L. Zagorchev and A. Goshtasby, A comparative study of transformation functions for nonrigid image registration, *IEEE Transactions on Image Processing*, **15**(3):529–538, 2006.

17 B. Zitova and J. Flusser, Image registration methods: a survey, *Image and Vision Computing*, **21**(11):977–1000, 2003.

18 L. G. Brown, A survey of image registration techniques, *ACM Computing Surveys*, **24**:326–376, 1992.

19 B. Dawant, G. E. Christensen, J. M. Fitzpatrick, and D. Rueckert, *Biomedical Image Registration: 5th International Workshop*, Springer, 2012.

20 J. C. Gee, J. B. Maintz, and M. W. Vannier, *Biomedical Image Registration: 2nd International Workshop*, Springer, 2003.

21 A. Goshtasby and J. Le Moigne, Image registration: an introduction, *Pattern Recognition*, **32**(1):1–2, 1999.

22 A. Goshtasby, L. Staib, C. Studholme, and D. Terzopoulos, Nonrigid image registration, *Computer Vision and Image Understanding*, **89**(2/3):109–113, 2003.

23 A. Goshtasby, *Image Registration: Principles, Tools, and Methods*, Springer, 2012.

24 D. L. G. Hajnal, J. V. Hill, and D. J. Hawkes, *Medical Image Registration*, CRC Press, 2001.

25 J. Le Moigne, N. S. Netanyahu, and R. D. Eastman, *Image Registration for Remote Sensing*, Cambridge University Press, 2012.

26 J. Modersitzki, *Numerical Methods for Image Registration*, Oxford University Press, 2003.

27 J. Modersitzki, *FAIR: Flexible Algorithms for Image Registration*, SIAM, 2009.

28 K. Rohr, *Landmark-Based Image Analysis: Using Geometric and Intensity Models*, Kluwer Academic, Boston, MA, 2001.

2

Image Orientation Detection

2.1 Introduction

Images that are in different orientations are the most difficult to register. Just as we can recognize objects in their upright poses more easily than in upside-down or rotated poses, computer algorithms can more easily locate patches in one image in another if the images are not rotated with respect to each other. To simplify the correspondence process, images have to be brought to the same orientation.

Unless an image contains a perfectly symmetric pattern, it will have a preferred orientation. Images of natural scenes often possess this property. If the preferred orientation of each image can be reliably determined, the images can be brought to the same orientation, simplifying the correspondence process. In the remainder of this chapter, preferred orientation will be referred to as *orientation*.

The orientation of an image depends on the geometric layout of the pattern within the image. If the images to be registered do not have a sufficient overlap, they may have different orientations. Images of a scene captured by different sensors may also have different orientations because the same scene can appear differently in the images. However, when the images are from the same sensor and there is significant overlap between the images, by knowing the orientations of the images, it should be possible to bring both images to the same orientation.

An image's (global) orientation is calculated from the aggregate of its local orientations, and a local orientation is determined from intensity gradients or geometric gradients in a small neighborhood. Since orientation is a geometric property, we will see that geometric gradients provide a more reliable means of estimating it than intensity gradients.

Radiometric and geometric noise can influence estimation of an image's orientation. Radiometric noise is usually zero-mean and can be reduced by image

Theory and Applications of Image Registration, First Edition. Arthur Ardeshir Goshtasby.

smoothing. However, geometric noise is not zero-mean, and if not removed, it can adversely influence estimation of an image's orientation.

To compare image orientations computed from intensity gradients and geometric gradients, consider the noise-free image in Fig. 2.1a. The image contains a solid circle with its radius taken in such a way that there would be 360 pixels on the boundary of the circle. Ideally, pixels along the boundary of the circle should produce uniform gradient directions between 0° and 359°. The histogram of intensity gradient directions of pixels with nonzero gradient magnitudes are shown in Fig. 2.1c in red. While there are many pixels with gradient directions that are a multiple of 45°, there are not any pixels at many other gradient directions.

Denoting the intensity at pixel (x, y) in image I by $I(x, y)$, intensity gradient in x-direction will be $I_x(x, y) = I(x + 1, y) - I(x, y)$, intensity gradient in y-direction will be $I_y(x, y) = I(x, y + 1) - I(x, y)$, intensity gradient magnitude $GM(x, y)$ will be

$$GM(x, y) = \sqrt{I_x^2(x, y) + I_y^2(x, y)} \tag{2.1}$$

and intensity gradient direction $\theta(x, y)$ will be

$$\theta(x, y) = \arctan\left(\frac{I_y(x, y)}{I_x(x, y)}\right). \tag{2.2}$$

By smoothing the image in Fig. 2.1a with a Gaussian filter of standard deviation 3 pixels and calculating the gradient directions at pixels with gradient magnitudes greater than 0.5, the blue histogram of Fig. 2.1c is obtained. A larger number of pixels belong to the blue histogram compared to the red histogram. Although there are no longer strong preferences toward $45°, 135°, 225°$, and $315°$ orientations, there are still strong preferences toward orientations that are a multiple of 90°.

By increasing the radius of the circle by a factor of 2 so that there will be 720 pixels along the boundary of the circle as shown in Fig. 2.1b, the gradient direction histograms shown in Fig. 2.1d are obtained when using the original image (red) and the image after smoothing with a Gaussian filter of standard deviation 3 pixels (blue). The preference of gradient directions toward angles that are a multiple of 90° remains, making global orientation detection from the aggregate of local intensity gradient directions unreliable, even in the absence of radiometric noise. Local orientations determined by intensity gradients are biased toward the horizontal and vertical arrangement of pixels in an image—something that cannot be avoided irrespective of the level of intensity smoothing applied to the image.

A second example demonstrating the unreliable nature of intensity gradients in the estimation of an image's orientation is given in Fig. 2.2. The major axis of the ellipse in Fig. 2.2a makes a 30° angle with the horizontal axis. Therefore,

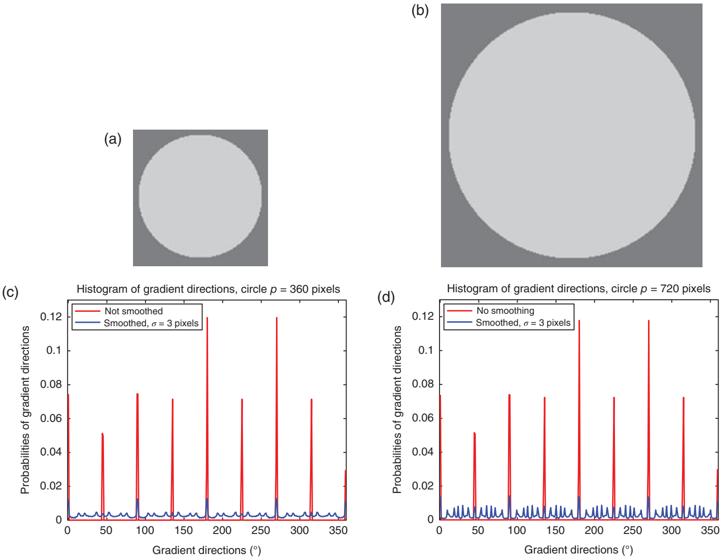
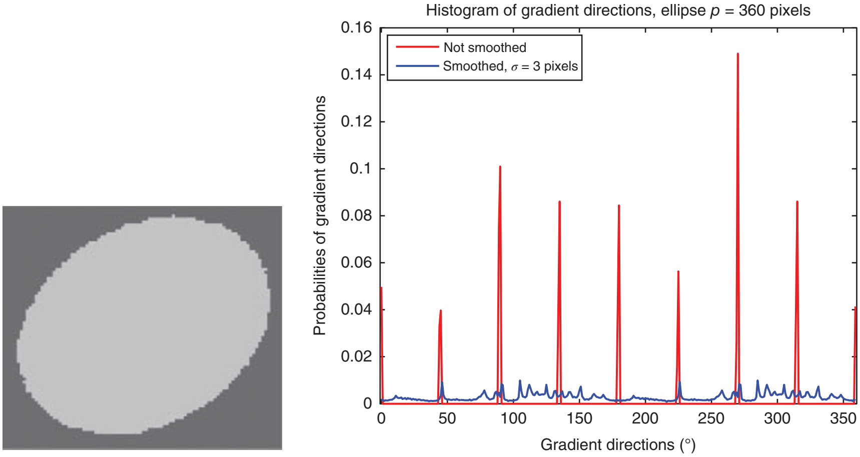


Figure 2.1 (a and b) Solid circles with perimeters 360 and 720 pixels, respectively. The intensity of pixels within the circles is 200, while the intensity of pixels outside the circles is 100. (c and d) Intensity gradient direction histograms of images (a) and (b), respectively, before smoothing (red) and after smoothing with a Gaussian of standard deviation 3 pixels (blue).



(a) (b)

Figure 2.2 (a) A solid ellipse with its major axis making a 30° angle with the *x*-axis. (b) Intensity gradient direction histograms of image (a) before smoothing (red) and after smoothing with a Gaussian of standard deviation 3 pixels (blue).

the preferred orientation of the image is clearly 30° or 210°. Nonetheless, when local orientations are calculated using intensity gradients, the histograms shown in Fig. 2.2b are obtained when using the image as is (red) and the image after smoothing with a Gaussian of standard deviation 3 pixels (blue). Before smoothing, strong preferences are observed toward angles that are a multiple of 45°. After smoothing, the preferences are not as strong but are still there. Considering that the intensity gradient direction at a boundary pixel is normal to the boundary contour, the peak intensity gradient direction in this image should be 120° or 300°.

The circle and the ellipse images in these examples are free of radiometric noise. When an image of a scene is captured, radiometric and geometric noises are both present. Radiometric noise is primarily sensor noise, and geometric noise is the displacement of scene points from their true positions after being projected to the discrete image space. Radiometric noise can be reduced by image smoothing; however, geometric noise will not disappear with intensity smoothing and will be present as long as the created image has discrete pixel coordinates.

To reliably determine the orientation of an image, two problems must be solved. First, the image should be converted into a form that is independent of absolute intensities and so insensitive to changes in scene lighting. Second, the digital domain of an image should be taken to a continuous one to remove/reduce geometric noise. To achieve this, the concepts of *geometric gradient* and *geometric smoothing* are introduced.

2.2 Geometric Gradient and Geometric Smoothing

To reduce the effect of sensor noise, an image is smoothed, and to make an image insensitive to changes in scene lighting, image edges rather than image intensities are used. Image edges can be detected by (1) the Canny edge detector [1], (2) zero-crossings of the second derivative intensities [2] followed by the removal of phantom edges [3, 4], (3) functional approximation [5], and (4) one-crossings of intensity ratios [6]. Edges obtained by these methods detect similar structures. Although Canny edge detector will be used below, other edge detectors may be used in the same manner. Canny edges represent pixels in an image where intensity gradient magnitude in the gradient direction becomes locally maximum. Canny edges appear in open and closed contours; they do not appear in structures with branches.

Edges are influenced by the horizontal and vertical arrangement of pixels in an image. To reduce this dependency, the digital domain should be converted into a continuous one by approximating each edge contour by a continuous parametric curve.

Given a sequence of pixels along an edge contour,

$$\mathbf{P} = \{\mathbf{p}_i = (x_i, y_i) \,:\, i = 0, \ldots, n-1\}, \tag{2.3}$$

the parametric curve approximating the contour takes the form

$$\mathbf{p}(u) = [x(u), y(u)], \tag{2.4}$$

where $x(u)$ and $y(u)$ are the components of the curve, defined in terms of the same parameter u. For the curve to approximate the contour, it is required that

$$\begin{aligned} x(u_i) &\approx x_i, \\ y(u_i) &\approx y_i, \end{aligned} \qquad \text{for } \; i = 0, \ldots, n-1. \tag{2.5}$$

In the following, the rational Gaussian (RaG) curve formulation [7] is used for its ability to approximate a curve of a desired smoothness to a sequence of points. The RaG curve approximating the sequence of edge pixels in $\mathbf{P}$ is defined by

$$\mathbf{p}(u) = \sum_{i=0}^{n-1} \mathbf{p}_i g_i(u), \qquad u \in [0, n-1], \tag{2.6}$$

where

$$g_i(u) = \frac{G_i(u)}{\sum_{j=0}^{n-1} G_j(u)}, \qquad i = 0, \ldots, n-1 \tag{2.7}$$

are the basis functions of the curve, and

$$G_i(u) = \exp\left\{ -\frac{(u - u_i)^2}{2\sigma^2} \right\} \tag{2.8}$$

is a Gaussian of height 1 and standard deviation σ centered at parameter u_i. Parameters $\{u_i : i = 0, \dots, n-1\}$ are the nodes of the curve. These are the parameters at which the curve approximates the individual pixels along the edge contour. σ is the smoothness parameter. By increasing it, the curve becomes smoother, and by decreasing it, the curve reproduces more details along the contour.

The equations above are used for an open curve. For a closed curve, Eq. (2.8) is replaced with

$$G_i(u) = \sum_{j=-\infty}^{j=\infty} \exp\left\{-\frac{[u-(u_i+jn)]^2}{2\sigma^2}\right\}. \tag{2.9}$$

Since a Gaussian vanishes exponentially from its center point, instead of ∞ in Eq. (2.9), a small number such as 1 or 2 is sufficient. The addition of jn to u_i or subtraction of jn from u_i ensures that at the point where the curve closes, the Gaussians centered at the nodes corresponding to the contour pixels extend over the closing point, producing a continuous and smooth curve everywhere.

Examples of curve fitting in this manner are given in Fig. 2.3. The Canny edges of the circle in Fig. 2.1a and the ellipse in Fig. 2.2a obtained with a Gaussian smoother of $\sigma = 3$ pixels are shown in Fig. 2.3a and b, respectively. The RaG curves approximating the contours are drawn in red over the edge contours.

As σ is increased, the approximating curve smoothes more details along the edge contour, and as σ is decreased, the curve follows the edge contour more closely, reproducing the details. In the following, the σ used in curve fitting is set to 3 pixels or higher. A σ lower than 3 pixels will reproduce geometric noise by closely following pixels along the contour and, therefore, is not recommended.

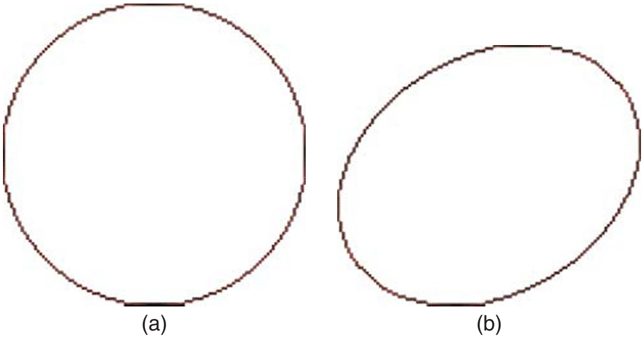


Figure 2.3 (a and b) RaG curves (shown in red) approximating the boundary contours of the solid circle in Fig. 2.1a and the solid ellipse in Fig. 2.2a obtained by the Canny edge detector when using a Gaussian smoother of standard deviation 3 pixels.

2.2.1 Calculating Geometric Gradients

The tangent direction θ at a curve point with parameter u is defined by

$$\theta(u) = \arctan\left(\frac{dy(u)}{dx(u)}\right) \tag{2.10}$$

or

$$\theta(u) = \arctan\left(\frac{dy(u)/du}{dx(u)/du}\right). \tag{2.11}$$

Since

$$\frac{dx(u)}{du} = \left.\frac{x(u+\delta u) - x(u)}{\delta u}\right|_{\delta u \to 0}, \tag{2.12}$$

to estimate $dx(u)/du$, $x(u)$ is computed at two very close values of u and their difference is divided by the difference in their u values. That is,

$$\frac{dx(u)}{du} = \frac{x(u+\delta u) - x(u-\delta u)}{2\delta u}. \tag{2.13}$$

Similarly,

$$\frac{dy(u)}{du} = \frac{y(u+\delta u) - y(u-\delta u)}{2\delta u}. \tag{2.14}$$

δu is a very small increment, such as 0.01. With such a small increment in u, the difference between the actual and estimated tangent directions will be negligible considering that θ will be quantized to create the histogram of tangent directions. All tangent directions are changed between 0 and π so that independent of the direction a contour is traced, the same tangent direction is obtained at a point. *The tangent direction $\theta(u_i)$ at point $\mathbf{p}(u_i)$ on the curve approximating contour* **P** *is considered the geometric gradient direction at pixel* $\mathbf{p}_i$ on the contour.

By computing geometric gradient directions along the RaG curve that approximates the circular edge contour in Fig. 2.3a at uniform increments $du = 0.01$ and creating the histogram of the geometric gradient directions, the histogram depicted in red in Fig. 2.4 is obtained. The histogram is almost flat, showing no significant preferred orientation. Due to large linear segments along the contour with angles that are a multiple of 45°, small bumps are observed at those angles. The bumps will get smaller by increasing the smoothness parameter σ in curve fitting. The histogram of geometric gradient directions for the curve that approximates the elliptic edge contour in Fig. 2.3b is depicted in blue in Fig. 2.4. The peak orientation at 30° is obvious.

To determine the curvature value at a curve point, first, $d^2x(u)/du^2$ and $d^2y(u)/du^2$ are estimated from changes in $dx(u)$ and $dy(u)$ over a small change

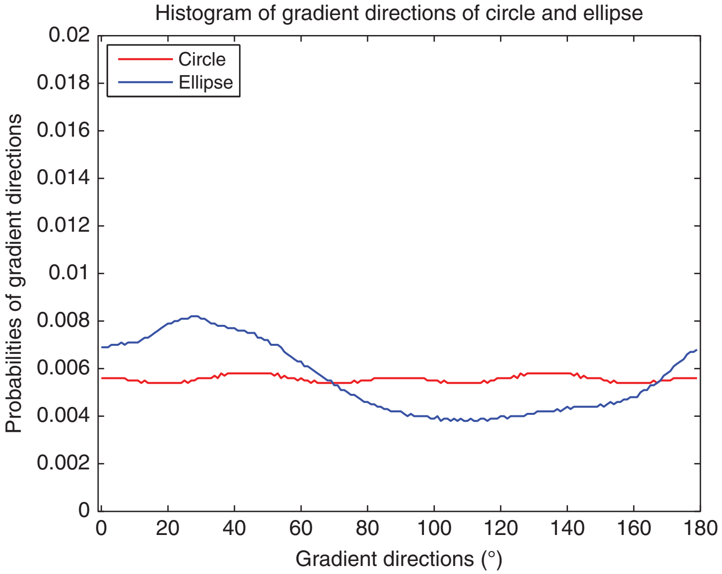


Figure 2.4 The red and blue plots show histograms of tangent directions along curves in Fig. 2.3a and b, respectively.

in u, that is,

$$\frac{d^2x(u)}{du^2} = \frac{dx(u+\delta u) - dx(u-\delta u)}{2\delta u}, \tag{2.15}$$

$$\frac{d^2y(u)}{du^2} = \frac{dy(u+\delta u) - dy(u-\delta u)}{2\delta u}, \tag{2.16}$$

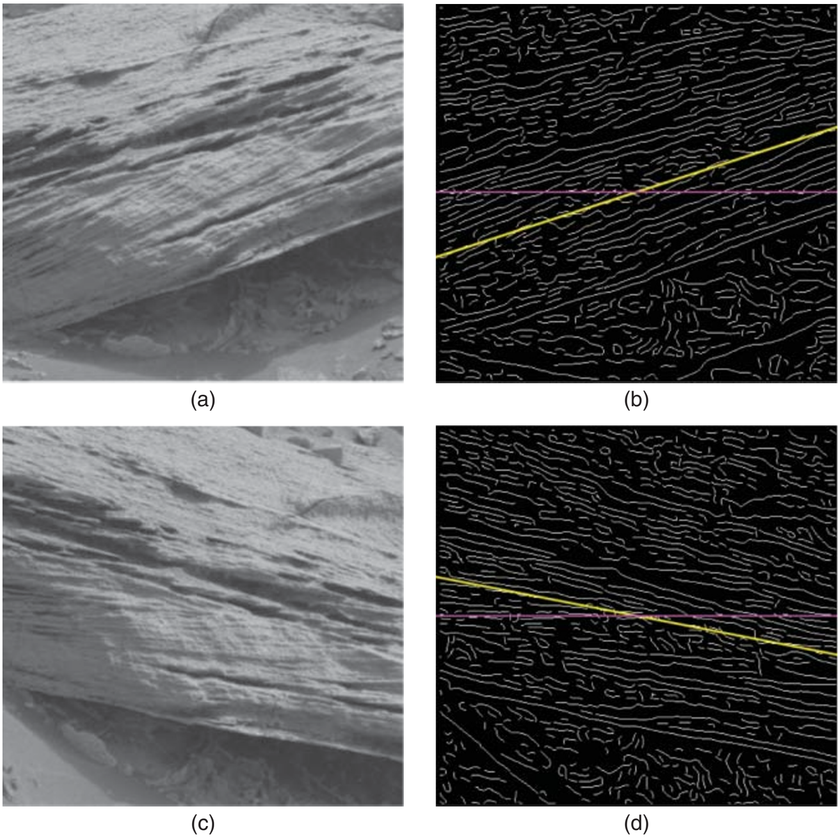
for $\delta u = 0.01$. Note that the difference between the parameters of curve points corresponding to adjacent pixels along a contour is 1. That is, $u_{i+1} - u_i = 1$ for $i = 0, \dots, n-2$.

Denoting $dx(u)/du, dy(u)/du, d^2x(u)/du^2$, and $d^2y(u)/du^2$ by x', y', x'', and y'', respectively, the magnitude curvature at the point with parameter u on the curve can be computed from [8]:

$$\kappa = \frac{|x'y'' - y'x''|}{(x'^2 + y'^2)^{3/2}}. \tag{2.17}$$

Since curvature is independent of the coordinate system of the curve or the edge contour it is approximating, it produces the same value at a curve point independent of the location and the orientation of the edge contour in an image.

Examples of image orientation detection by geometric gradients using real images are given in Fig. 2.5. Figure 2.5a shows a Martian rock, and (b) shows its Canny edges. The peak geometric gradient direction of RaG curves with smoothness parameter $\sigma = 3$ pixels fitting the edge contours is found to be 18°,



(a) (b)

(c) (d)

Figure 2.5 (a and c) Images of the same Martian rock with a known rotational difference of 30°. These images are courtesy of NASA. (b and d) Canny edges of images (a) and (c), respectively, obtained when using a Gaussian smoother of standard deviation 3 pixels. Image orientations determined from geometric gradients and intensity gradients are shown with yellow and purple lines, respectively.

shown by the yellow line in Fig. 2.5b. The peak intensity gradient direction for Fig. 2.5a after smoothing with a Gaussian filter of standard deviation 3 pixels is found to be 90°. Since the intensity gradient direction at a pixel along an edge contour is normal to the edge contour, the image orientation obtained by intensity gradient directions will be 0°, as depicted in Fig. 2.5b by the purple line. In order to compare image orientations determined by geometric gradients and intensity gradients, *in all examples given in this chapter, the direction normal to the intensity gradient direction at a pixel is used as the intensity gradient direction there.*

Figure 2.5a represents a small window of a much larger image. By rotating the image by 30° clockwise and taking a window of the same size from the same area, the image in Fig. 2.5c is obtained. Canny edges of this image are shown in Fig. 2.5d. The peak geometric gradient direction is found to be 169°, shown by the yellow line in this image. The peak intensity gradient direction is found to be 90° with the direction normal to it being 0°, shown by the purple line in Fig. 2.5d. The difference in orientations of the yellow lines in images (b) and (d) is $18^\circ - 169^\circ = -151^\circ \equiv 29^\circ$, because no distinction is made between θ and $\theta \pm \pi$. The true rotational difference between the images is 30°. The rotational difference between the images estimated by geometric gradients is off by only 1°.

Figure 2.6a and b shows the RaG curves approximating the edge contours in Fig. 2.5b and d, respectively. Curve points drawn with brighter red are points with higher curvatures. When calculating the histogram of geometric gradient directions, instead of incrementing the histogram bin representing direction θ by 1, the bin is incremented by $1 - \kappa$, where κ is the curvature at the point, after normalizing all curvatures to values between 0 and 1. The process assigns higher values to curve points with lower curvatures, emphasizing straight segments over curved ones. The histograms of geometric gradient directions obtained in this manner for Fig. 2.6a and b are shown in red and blue, respectively, in (c).

In the following experiments, the standard deviation of the Gaussian smoother in the Canny edge detector is set to 3 pixels and the smoothness parameter of RaG curves approximating the edge contours is set to 5 pixels. Also, the tangent direction at a curve point is weighted by $w = 1 - \kappa$, with κ being the curvature value at a curve point after normalizing all curvatures to values between 0 and 1.

2.3 Comparison of Geometric Gradients and Intensity Gradients

To compare the sensitivities of geometric gradients and intensity gradients to variations in intensities and geometry of an image, the intensities and geometry of an image are varied and the stabilities of image orientations determined by the two methods are compared.

Figure 2.7a depicts a Martian scene. This image is courtesy of NASA. The image does not have a strong orientation, but it contains a unique pattern; therefore, it must have a unique orientation. Figure 2.7b shows the image after smoothing with a Gaussian of standard deviation of 2 pixels. The image after addition of Gaussian noise of standard deviation 20 is shown in Fig. 2.7c. If by this addition intensity became less than 0, it was set to 0, and if it became greater than 255, it was set to 255. The image in Fig. 2.7d is obtained by transforming the intensities in Fig. 2.7a by the following nonlinear function:

$$O(x, y) = I(x, y) + 50 \sin(4\pi y/n_r) \cos(4\pi x/n_c). \tag{2.18}$$

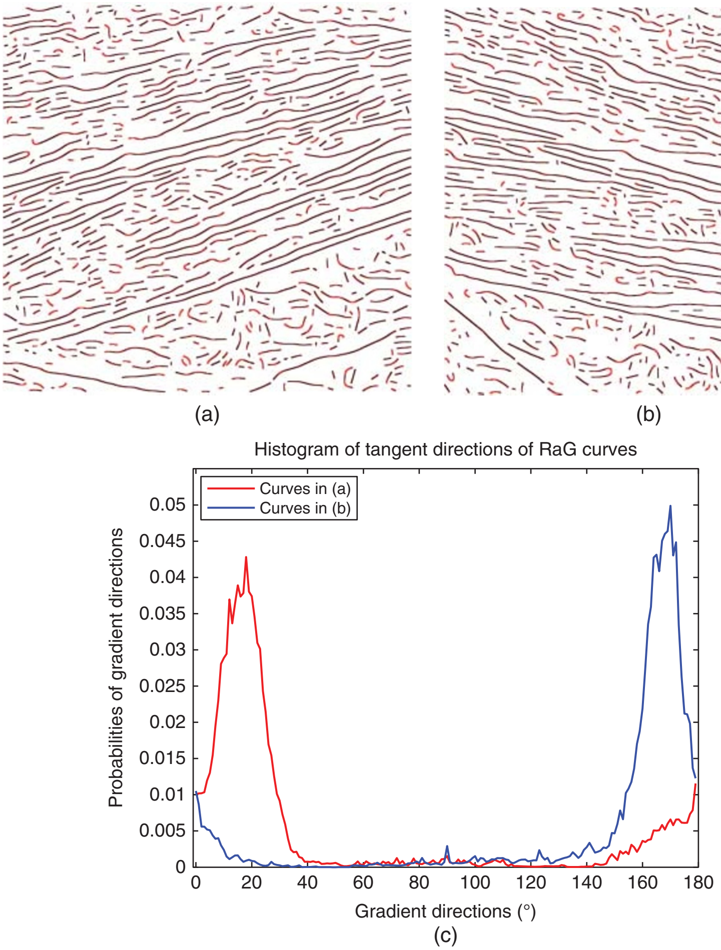
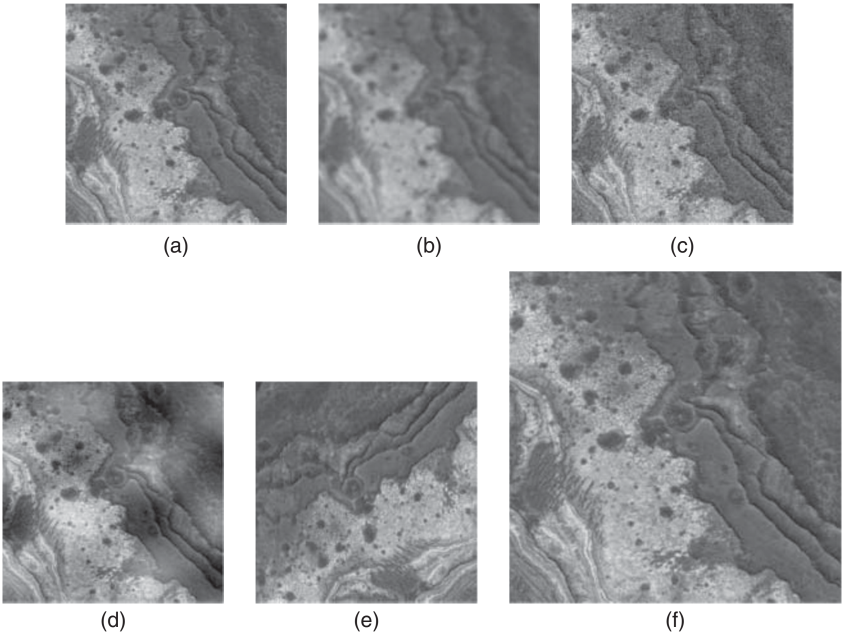


Figure 2.6 (a and b) RaG curves with $\sigma = 3$ approximating the edge contours in Fig. 2.5b and d, respectively. A point with a higher curvature is shown with a higher red intensity. (c) Red and blue plots show histograms of geometric gradient directions or tangent directions of points along curves in (a) and (b), respectively.

$I(x, y)$ is the intensity at (x, y) in Fig. 2.7a and $O(x, y)$ is the intensity at the same location in image (d). Parameters n_r and n_c show the number of rows and the number of columns in the images. Figure 2.7e is obtained by rotating image (a) by 90°. Finally, Fig. 2.7f is obtained by scaling image (a) by 1.5 using bilinear interpolation. Figure 2.7a–f contains the same pattern but have somewhat different intensities or have different orientations and scales. The ability of geometric gradients and intensity gradients in determining the orientations of these images is investigated.



(a) (b) (c) (d) (e) (f)

Figure 2.7 (a) An image of a Martian scene. This image is courtesy of NASA. (b) Blurred, (c) noisy, (d) nonlinear intensity mapping, (e) rotated, and (f) scaled versions of image.

Orientations of the images in Fig. 2.7 determined by geometric gradients and intensity gradients are shown by yellow and purple lines, respectively, in Fig. 2.8. The lines in each image pass through the image center and have the orientations obtained from the highest peaks in the histograms computed from geometric gradient directions and intensity gradient directions. The orientations estimated by geometric gradients and intensity gradients are listed in Table 2.1. While geometric gradients assign an orientation to an image that represents the orientation of the pattern, the orientation assigned to an image by intensity gradients is more representative of the arrangement of pixels in the image than the orientation of the pattern. Due to preferences toward angles that are a multiple of 45° by intensity gradients, when a pattern with a strong orientation does not appear in an image, the orientation assigned to the image will be a multiple of 45°.

Blurring, noise, and change in scene lighting hardly affect estimation of an image's orientation by geometric gradients. By performing computations in the continuous domain, geometric gradients make the orientation detection process independent of the horizontal and vertical arrangement of pixels in an image. Geometric gradients assign an orientation to the pattern in an image independent of the way the pattern is pixelated.

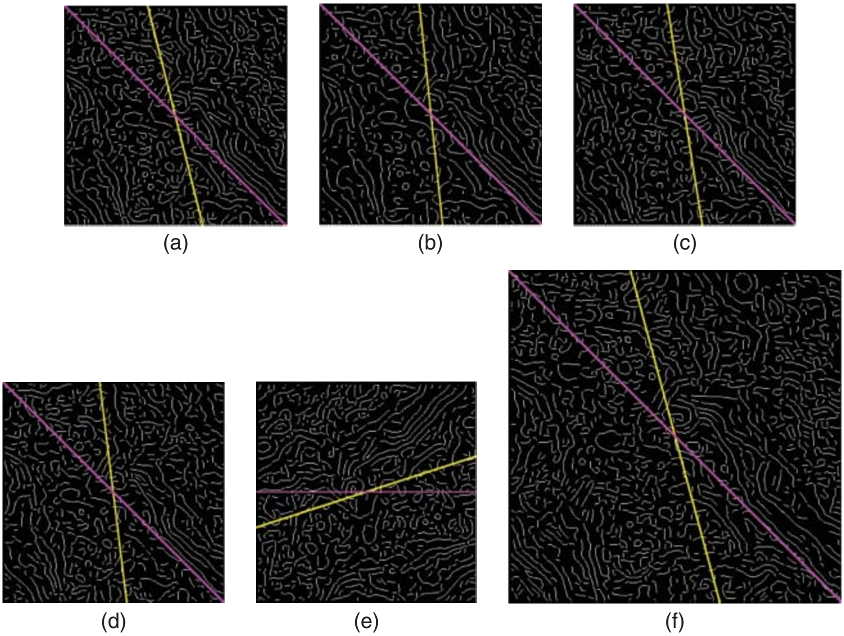


Figure 2.8 (a)–(f) Canny edges of images (a)–(f) in Fig. 2.7. Image orientations obtained by geometric gradients and intensity gradients are shown by yellow and purple lines, respectively.

Table 2.1 Orientations of images (a)–(f) in Fig. 2.8 estimated by geometric gradients (θ_{gg}) and intensity gradients (θ_{ig}).

Figure 2.8	θ_{gg}	θ_{ig}
(a)	104	135
(b)	96	135
(c)	99	135
(d)	97	135
(e)	16	0
(f)	105	135

2.4 Finding the Rotational Difference between Two Images

Now that we have a method for determining the orientation of an image, we can use it to determine the rotational difference between two images and rotate one image to the orientation of the other. Once the images are in the same orientation, the correspondence process becomes much easier, simplifying the registration process.

If the pattern appearing in two images has a strong preferred orientation, the difference in the peak gradient directions of the images is sufficient to find the rotational difference between the images. However, when the pattern appearing in the images does not have a strong orientation, information only about the peak orientations may not be sufficient to determine the rotational difference between the images, and information about all orientations may be needed to robustly find the rotational difference between the images.

Sometimes, the images to be registered may not contain exactly the same scene parts. A part of a scene visible in one image may be hidden in the other image due to occlusion or by falling outside the view. When the scene pattern does not contain a strong orientation or the overlap between the images is not large enough, the peak orientation determined in one image may not be the same peak orientation in the other image. By using all gradient directions rather than only the peak gradient direction in each image, the rotational difference between the images can be more reliably determined.

To compare two gradient direction histograms, one histogram is shifted cyclically over the other, and at each shift position, the Euclidean distance between the histograms is determined, and the shift position producing the smallest Euclidean distance is used as the rotational difference between the images. Govindu and Shekhar [9] show that the shift amount producing the smallest Euclidean distance between two histograms provides the maximum likelihood estimation of the translational difference between the histograms and, thus, shows the maximum likelihood rotational difference between the images the histograms are computed from.

Denoting the gradient direction histograms of two images by $H_1(\theta)$ and $H_2(\theta)$, the Euclidean distance between the histograms when the second histogram is cyclically shifted over the first by θ_i is

$$D(\theta_i) = \left\{ \sum_{j=0}^{m-1} [H_1(\theta_j) - H_2(\theta_j + \theta_i)]^2 \right\}^{\frac{1}{2}}, \tag{2.19}$$

where m is the number of bins in each histogram. In our case, $m = 180$. By varying i from 0 to 179 and computing the Euclidean distance D, the shift value producing the smallest distance measure is determined and used as the rotational difference between the images. Note that cyclic shifting of H_2 with respect to H_1 implies that when $\theta_j + \theta_i \geq m$, $H_2(\theta_j + \theta_i - m)$ is taken as the value for $H_2(\theta_j + \theta_i)$.

Although discrete angles may be acceptable when determining the orientation of an image, it may not be sufficient when determining the rotational difference between two very large images. An error of a single degree can result in displacement of a few to several pixels when registering the images.

To determine the subdegree rotational difference between two images, assuming the minimum distance between two histograms when using the Euclidean norm is θ, we fit a quadratic curve to $[(\theta - 1), D(\theta - 1)]$, $[\theta, D(\theta)]$, $[(\theta + 1), D(\theta + 1)]$, find the location of the minimum in the curve, and use that as the rotational difference between the images with subdegree accuracy.

The histograms of geometric gradient directions of images (a)–(f) in Fig. 2.8 are shown in Fig. 2.9. Computing the subdegree rotational difference between histogram (a) and histogram (b) by shifting histogram (b) over histogram (a) and finding the shift amount where the Euclidean distance between the histograms becomes minimum, we find the rotational difference of image (b) with respect to image (a) be 179.7° or −0.3°. Similarly, we find image (c) to be rotated with respect to image (a) by 0.2°, image (d) to be rotated with respect to image (a) by 0.1°, image (e) to be rotated with respect to image (a) by 90.0°, and image (f) to be rotated with respect to image (a) by 179.7° or −0.3°. Use of all bins in histograms of geometric gradient directions provides a more robust means of determining the rotational difference between two images when compared to the use of only histogram bins corresponding to the peak orientations.

2.5 Performance Evaluation

In this section, reliability, accuracy, and computational complexity of geometric gradients in image orientation detection are evaluated.

2.5.1 Reliability

To ascertain the ability of geometric gradients in determining the orientation of various types of images, experiments were carried out using aerial images (Fig. 2.10), close-range images captured from different camera views (Fig. 2.11), images with perspective differences (Fig. 2.12), images with radiometric as well as structural differences (Fig. 2.13), images with scale and rotational differences (Fig. 2.14), and images from different sensors (Fig. 2.15).

Edges detected in the images in Figs 2.10–2.15 are shown in Figs 2.16–2.21, respectively. Edges in color images are obtained by the method outlined in [6, 10]. Peak tangent directions obtained from the edges in images (a) and (b) in Fig. 2.16 are 15° and 19°, drawn with yellow lines in the images. By matching the gradient direction histograms of the images, image (b) is found to be rotated with respect to image (a) by 3.2°. The preferred orientations determined for these images by intensity gradients were both 0°, shown by purple lines in the images. Due to the horizontal and vertical preferences of intensity gradients, direction 0° has been selected for both images.

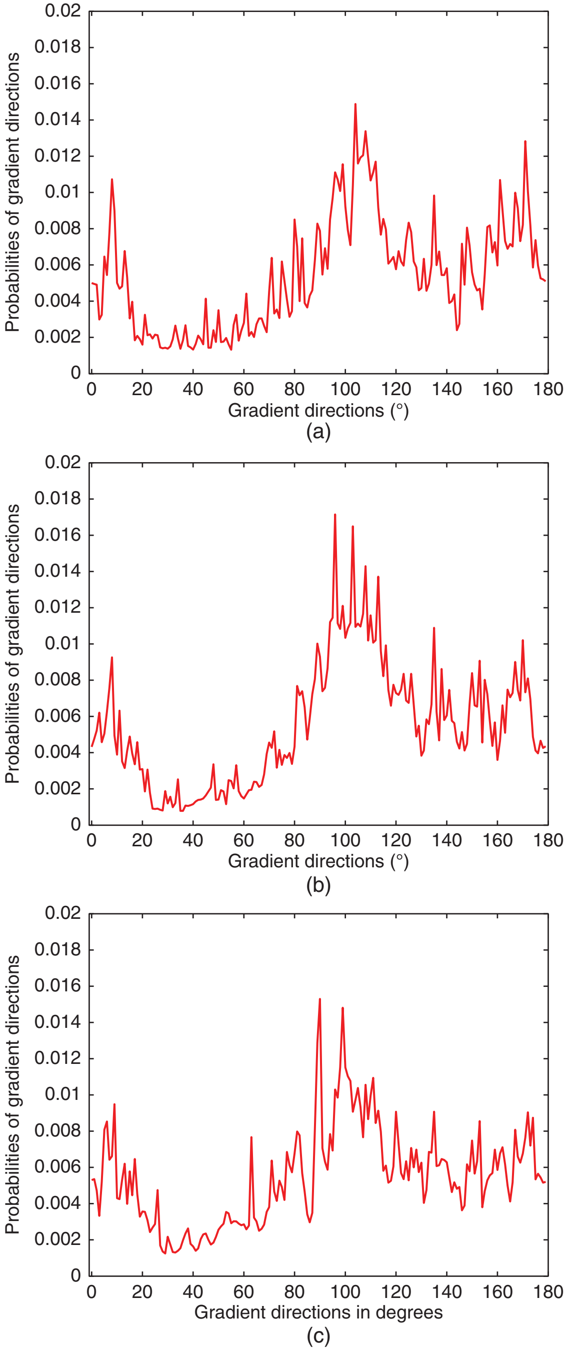
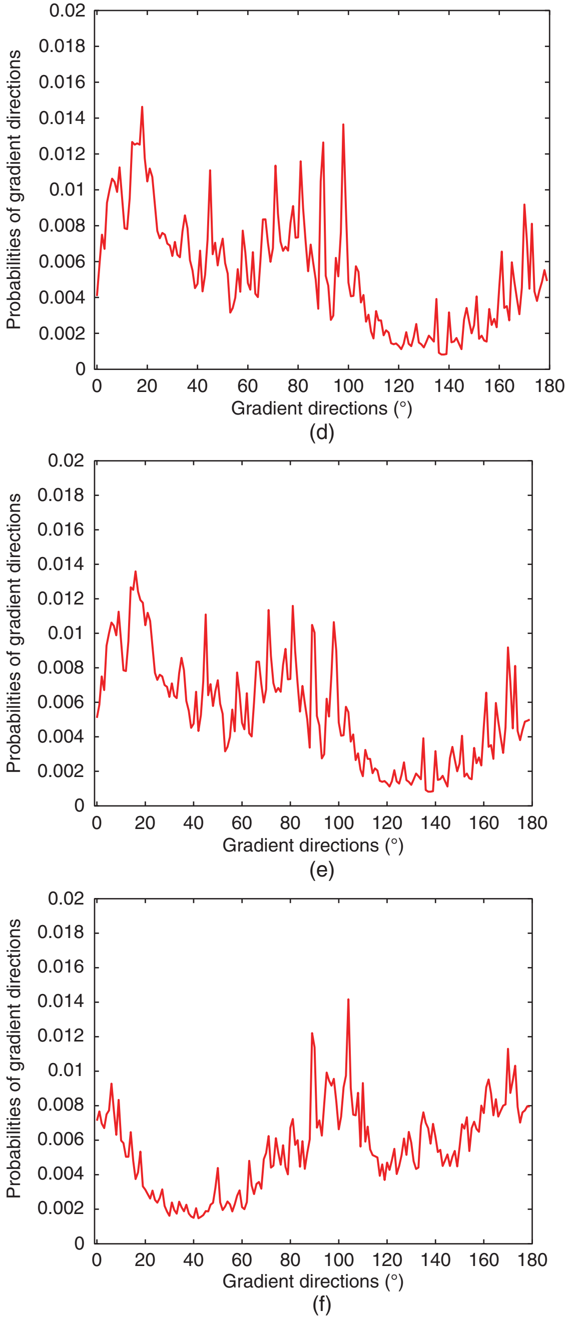


Figure 2.9 (a)–(f) Histograms of geometric gradient directions of images (a)–(f) in Fig. 2.8.

Figure 2.9 (*Continued*)



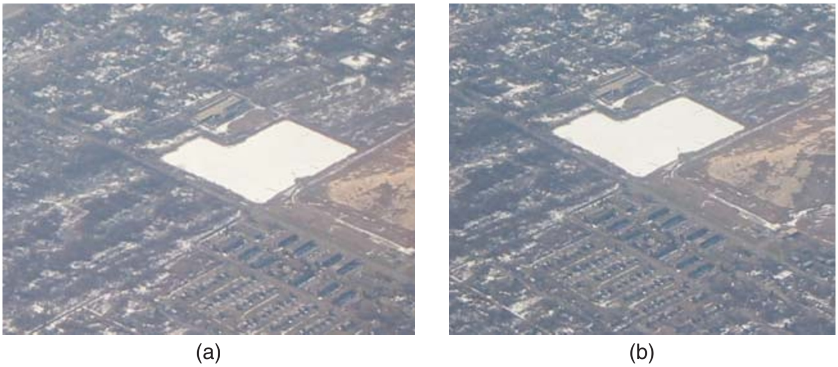

(a) (b)

Figure 2.10 (a and b) Aerial images taken from different views of a suburban area.

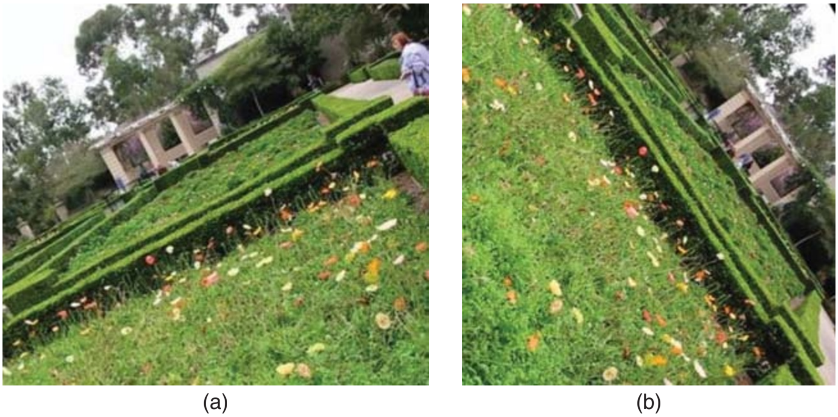

(a) (b)

Figure 2.11 (a and b) Images of a garden scene captured from different views.

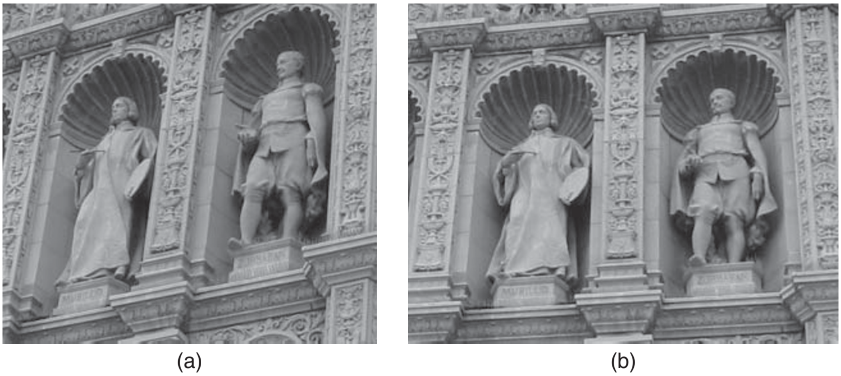
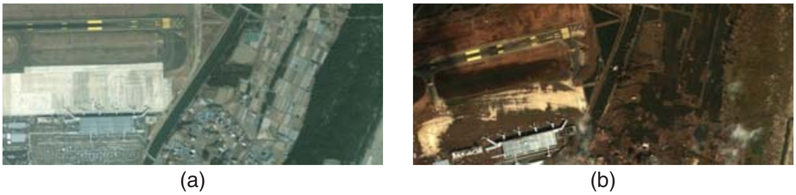
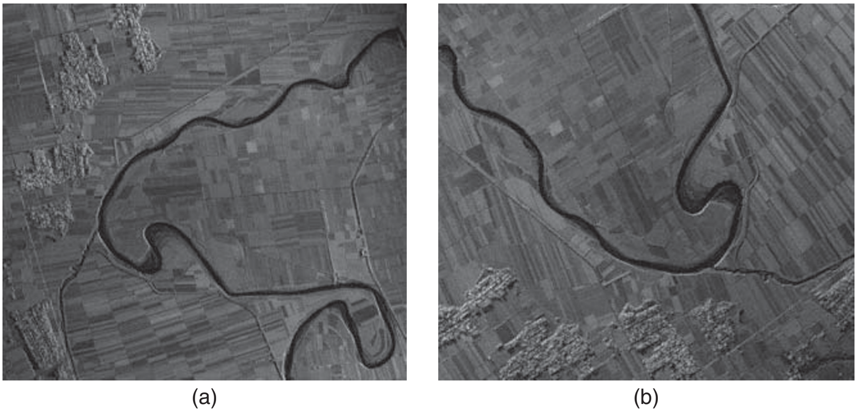

(a) (b)

Figure 2.12 (a and b) Images captured of a historical building from different views.



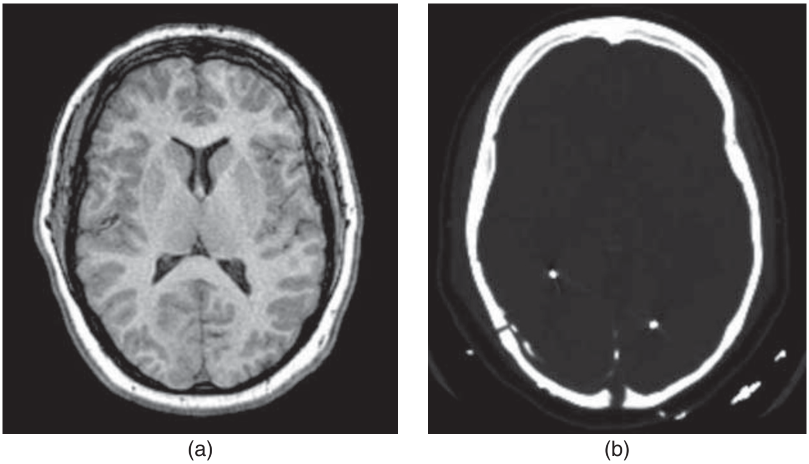
(a) (b)

Figure 2.13 (a and b) Aerial images of the Sendai International Airport taken before and after the grand tsunami of 2011. These images are courtesy of Google Maps.



(a) (b)

Figure 2.14 (a and b) Radar images of a relatively flat agricultural area captured in different flight paths. Image (b) is artificially scaled to introduce scale difference between images. These images are courtesy of Radarsat-1.



(a) (b)

Figure 2.15 (a and b) Corresponding axial slices from registered MR and CT brain volumes. These images are courtesy of Kettering Medical Center.

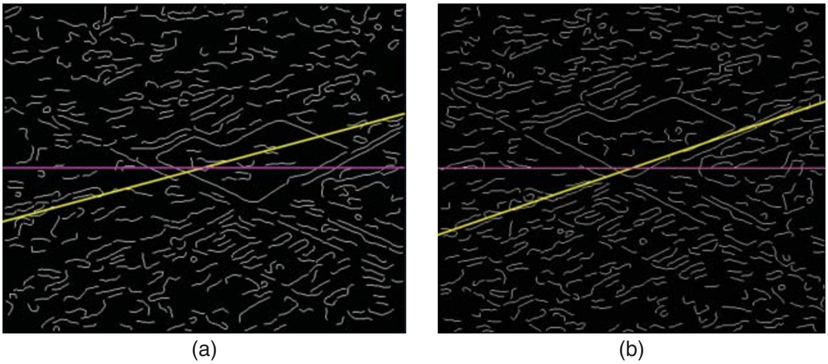


Figure 2.16 (a and b) Canny edges of aerial images (a) and (b) in Fig. 2.10 and the orientations detected by geometric gradients (yellow) and intensity gradients (purple).

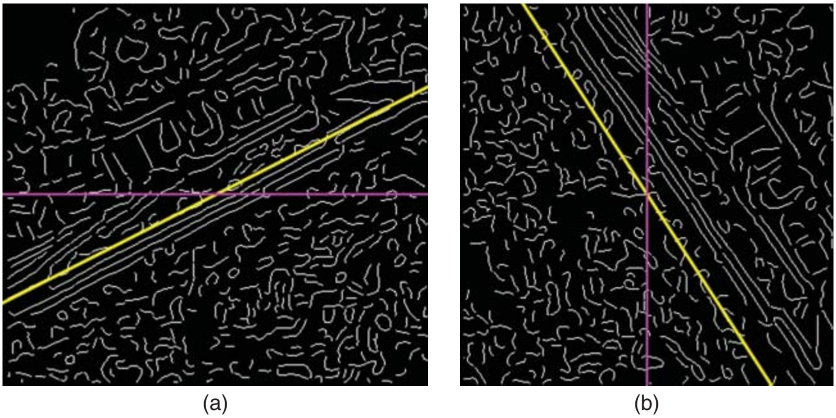


Figure 2.17 (a and b) Canny edges of the garden scene images (a) and (b) in Fig. 2.11. The orientations detected for these images by geometric gradients and intensity gradients are shown with yellow and purple lines, respectively.

The peak orientations determined for the garden scene images in Fig. 2.11 are shown in Fig. 2.17. The peak orientations determined by geometric gradients are 27° and 123° for images (a) and (b), respectively. These peak orientations are shown with yellow lines in the images. By matching the histograms of geometric gradient directions, image (b) is found to be rotated with respect to image (a) by 96.4°. The peak orientations determined from the histograms of intensity gradient directions in images (a) and (b) are found to be 0° and 90°, respectively. Peak orientations determined by the intensity gradients are shown with purple lines in these images.

Large perspective differences exist between images (a) and (b) in Fig. 2.12. The orientations determined from geometric gradients of the edge images

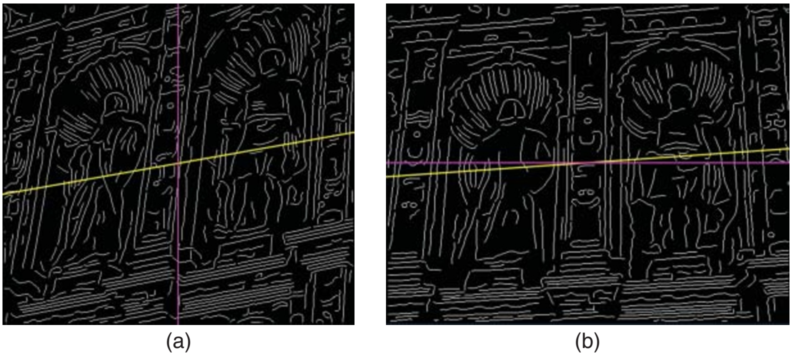

(a) (b)

Figure 2.18 (a and b) Canny edges of the historical building images (a) and (b) in Fig. 2.12. The orientations detected by geometric gradients and intensity gradients are shown with yellow and purple lines, respectively.

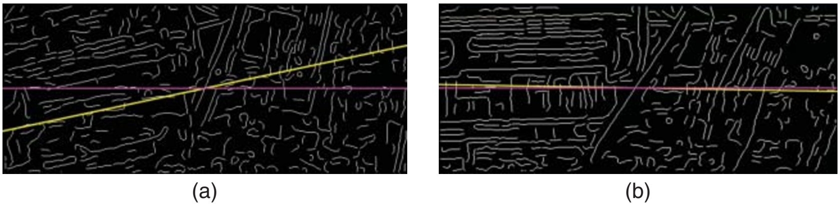

(a) (b)

Figure 2.19 (a and b) Canny edges of the before and after tsunami images (a) and (b) in Fig. 2.13. The orientations detected by geometric gradients and intensity gradients are shown by yellow and purple lines, respectively.

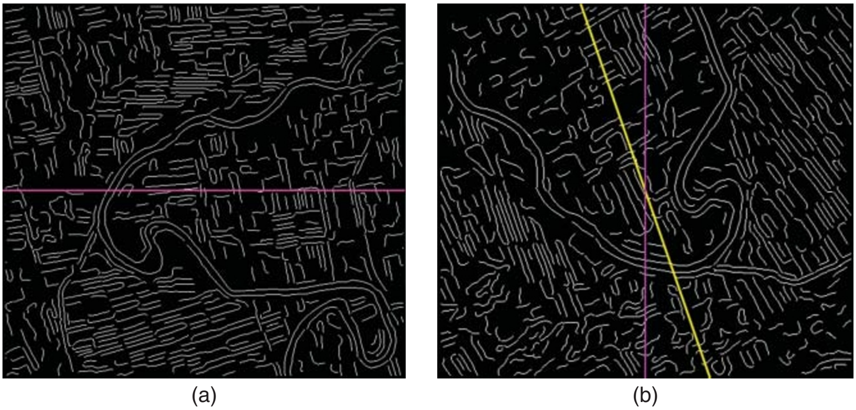

(a) (b)

Figure 2.20 (a and b) Canny edges of the radar images (a) and (b) in Fig. 2.14. The orientations detected by geometric gradients and intensity gradients are shown with yellow and purple lines, respectively. The yellow line in (a) is hidden by the purple line and, thus, is not visible.

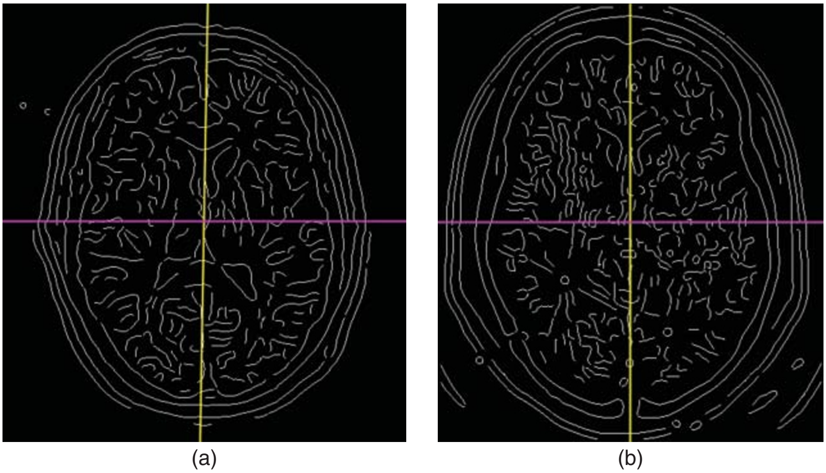


Figure 2.21 (a and b) Canny edges of the MR and CT brain slices in Fig. 2.15. The orientations determined by geometric gradients and intensity gradients are shown with yellow and purple lines, respectively.

depicted in Fig. 2.18 are 10° and 4°, respectively. By matching the histograms of geometric gradient directions, image (b) is found to be rotated with respect to image (a) by 173.5° or −6.5°. The orientations determined from the histograms of intensity gradient directions are 90° and 0°, shown with purple lines in Fig. 2.18.

Figure 2.13a and b not only have very different colors due to change in ground cover, there are also considerable structural changes in the scene caused by destructions from tsunami. Examining the images we see, however, that certain scene structures remain unchanged. Such structures are the airport's runway, the nearby streets, and waterways. Peak orientations obtained from edge images (a) and (b) in Fig. 2.19 are 12° and 179°, respectively. The process has assigned the same orientation to the scene independent of the image differences. Through histogram matching, image (b) is found to be rotated with respect to image (a) by 13.2°. The orientations computed from intensity gradient directions are found to be both 0°, shown by purple lines in the images.

Figure 2.14 has rotational and scaling differences. The geometric gradient direction peaks calculated from the edge images in Fig. 2.20 are 0° and 109°. These orientations are shown with yellow lines in the images. By matching the gradient direction histograms of the images, image (b) is found to be rotated with respect to image (a) by 109.2°. The orientations determined from the peaks in the intensity gradient direction histograms of the images are 0° and 90°. These orientations are shown with purple lines in the images. 0° orientation is detected by both methods for image (a), resulting in the yellow and purple