Intelligent Robots & Computer Vision VIII: Algorithms & Techniques

PROCEEDINGS
SPIE—The International Society for Optical Engineering

Intelligent Robots and Computer Vision VIII: Algorithms and Techniques

David P. Casasent
Chair/Editor

6–10 November 1989
Philadelphia, Pennsylvania

Sponsored by
SPIE—The International Society for Optical Engineering

Cooperating Organizations
Center for Optical Data Processing/Carnegie Mellon University
IEEE Philadelphia Section
The Industrial Electronics Society of the IEEE

Published by
SPIE—The International Society for Optical Engineering
P.O. Box 10, Bellingham, Washington 98227-0010 USA
Telephone 206/676-3290 (Pacific Time) • Telex 46-7053

P

Volume 1192
Part One of Two Parts

The papers appearing in this book comprise the proceedings of the meeting mentioned on the cover and title page. They reflect the authors' opinions and are published as presented and without change, in the interests of timely dissemination. Their inclusion in this publication does not necessarily constitute endorsement by the editors or by SPIE.

Please use the following format to cite material from this book:
Author(s), "Title of Paper," *Intelligent Robots and Computer Vision VIII: Algorithms and Techniques,* David P. Casasent, Editor, Proc. SPIE 1192, page numbers (1990).

Library of Congress Catalog Card No. 89-63974
ISBN 0-8194-0231-1

Address inquiries and notices to Director of Publications, SPIE, P.O. Box 10, Bellingham, WA 98227-0010 USA.

INTELLIGENT ROBOTS AND COMPUTER VISION VIII:
ALGORITHMS AND TECHNIQUES

Volume 1192

CONFERENCE COMMITTEE

Chair
David P. Casasent, Carnegie Mellon University

Cochairs
Ernest L. Hall, University of Cincinnati

Program Committee
Rolf-Jürgen Ahlers, Fraunhofer-Institut für Produktionstechnik und Automatisierung (FRG)
Bruce G. Batchelor, University of Wales (UK)
Mark F. Cullen, Perkin-Elmer Corporation
Volker Gerbig, Siemens Erlanger (FRG)
Madan M. Gupta, University of Saskatchewan (Canada)
James M. Keller, University of Missouri/Columbia
Richard A. Messner, University of New Hampshire
Steven K. Rogers, Air Force Institute of Technology
Matthew A. Turk, Massachusetts Institute of Technology
William E. Wolf, E. I. du Pont de Nemours & Company, Inc.

Session Chairs
Session 1—Pattern Recognition for Intelligent Robots and Computer Vision
David P. Casasent, Carnegie Mellon University

Session 2—Segmentation, Image Processing, and Feature Extraction
Frederick M. Waltz, 3M Company

Session 3—Three-Dimensional Shape Determination and Representation
William E. Wolf, E. I. du Pont de Nemours & Company, Inc.

Session 4—Color and Range Image Processing
Matthew A. Turk, Massachusetts Institute of Technology

Session 5—Neural Networks and Associative Processors for Advanced Vision Processing
Steven K. Rogers, Air Force Institute of Technology

Session 6—Biological Basis for Machine Vision
Madan M. Gupta, University of Saskatchewan (Canada)

Session 7—Fuzzy Logic in Intelligent Systems and Computer Vision
James M. Keller, University of Missouri/Columbia

INTELLIGENT ROBOTS AND COMPUTER VISION VIII:
ALGORITHMS AND TECHNIQUES

Volume 1192

Session 8—Image Understanding and Analysis
Mark F. Cullen, Perkin-Elmer Corporation

Session 9—Time-Sequential Image Processing
Wang-he Lou, Cornell University

Session 10—Polar Exponential Grid Processing for Synthetic Vision Systems
Richard A. Messner, University of New Hampshire

INTELLIGENT ROBOTS AND COMPUTER VISION VIII:
ALGORITHMS AND TECHNIQUES

Volume 1192

INTRODUCTION

This eighth conference in this series continues the tradition of presenting the newest research results and developments in intelligent robots and computer vision. Emphasis is given to new techniques and algorithms.

More than 80 papers from ten countries were included in ten sessions that address various new advances and new thrusts in this area. The papers in this volume are arranged into topical sessions on specific approaches and problems in machine vision. Session 1, on pattern recognition, included new algorithms for occluded and distortion-invariant object identification, computer-vision progress on recognition by components, and a real-time laboratory system for locating lines in a product to be inspected. Session 2, on image processing, contained new advances in segmentation, edge detection, and feature extraction. Session 3, on three-dimensional shape determination and representation, and Session 4, on range image processing, provided new algorithms and results for this formidable problem. Session 5, on neural networks, represented advanced techniques for automating analyses of difficult computer-vision and robotic problems. Session 7, on fuzzy logic, and Session 8, on image understanding and analysis, addressed yet other techniques and new algorithms and results in this area. New sessions were included on biologically motivated machine vision systems (Session 6), time-sequential image processing of multiple frames of data (Session 9), and polar exponential grid processing techniques and hardware for in-plane distortion invariance (Session 10).

This conference has spawned various other conferences on more specific topics that should be of interest to many readers. These include

- Intelligent Robots and Computer Vision VIII: Systems and Applications (SPIE Volume 1193),
- Optics, Illumination, and Image Sensing for Machine Vision IV (SPIE Volume 1194),
- Mobile Robots IV (SPIE Volume 1195),
- Intelligent Control and Adaptive Systems (SPIE Volume 1196),
- Automated Inspection and High Speed Vision Architectures III (SPIE Volume 1197), and
- Sensor Fusion II: Human and Machine Strategies (SPIE Volume 1198).

I thank my secretary Marlene Layton, my cochair, the program committee, all session chairs, and the authors who made this conference the success it was.

David P. Casasent
Carnegie Mellon University

INTELLIGENT ROBOTS AND COMPUTER VISION VIII:
ALGORITHMS AND TECHNIQUES

Volume 1192

CONTENTS

(continued)

INTELLIGENT ROBOTS AND COMPUTER VISION VIII:
ALGORITHMS AND TECHNIQUES

Volume 1192

INTELLIGENT ROBOTS AND COMPUTER VISION VIII:
ALGORITHMS AND TECHNIQUES

Volume 1192

Part Two

(continued)

INTELLIGENT ROBOTS AND COMPUTER VISION VIII:
ALGORITHMS AND TECHNIQUES

Volume 1192

INTELLIGENT ROBOTS AND COMPUTER VISION VIII:
ALGORITHMS AND TECHNIQUES

Volume 1192

INTELLIGENT ROBOTS AND COMPUTER VISION VIII:
ALGORITHMS AND TECHNIQUES

Volume 1192

SESSION 1

Pattern Recognition for Intelligent Robots and Computer Vision

Chair
David P. Casasent
Carnegie Mellon University

Real-Time Optical Hough Transform for Industrial Inspection

Jeffrey Richards and David Casasent
Carnegie Mellon University
Department of Electrical and Computer Engineering
Center for Excellence in Optical Data Processing
Pittsburgh, PA 15213

Abstract

We describe a real-time optical Hough transform (HT) inspection system and show quantitative inspection results using an industrial inspection application. The HT architecture uses an electronically addressed liquid crystal television (LCTV) as the real-time spatial light modulator, a novel selective edge-enhancement filtering technique, and realizes multiple slices of the HT with a computer generated hologram. The industrial case study of the inspection of cigarette packages is used to benchmark the HT processor. A test set of 100 packages is presented to the processor to qualify its effectiveness. The statistical significance of these finite test set results is also examined.

1 Introduction

With the increasing industrial emphasis on quality control, inspection is quickly becoming an important criterion in many factory processes. Typically, commercially available digital machine vision systems are used for such inspections. Optical processing techniques have the potential to implement many of the same inspection algorithms as digital systems with increased speed and accuracy, and at lower costs. An optical processing vision system (based on Fourier transform (FT) wedge and ring features) has recently been introduced as a commercial product.[1,2,3] Other general purpose optical inspection processors for industry based on Fourier features have also been described.[4] Fourier features are often used in optical inspection applications because optics can easily generate the Fourier features[5]. However, many applications cannot be solved with Fourier features and more advanced algorithms and techniques are necessary. We have extended optical inspection techniques to include the Hough transform (HT) and show how the HT is used for inspection and can be performed optically at high speed with an inexpensive architecture. We will illustrate our real-time HT optical inspection processor on the industrial application of the inspection of cigarette packages.

Section 2 describes the optical architecture used in our optical HT inspection processor. Section 3 describes the industrial inspection case study of cigarette package inspection which we will use to benchmark our inspection system. Section 4 details the parameters of the optical laboratory system. Section 5 summarizes the results of the optical cigarette package inspections and analyzes the statistical significance of these results. Section 6 presents a summary and our conclusions.

2 HT Inspection Architecture

The computer generated hologram (CGH) has become increasingly popular in the past few years as a versatile optical element[6,7]. We have proposed[8] a new CGH which generates multiple slices of the HT in

parallel. The HT is useful as a feature space in inspection applications since the significant information in products being inspected is often straight lines (which the HT transforms to peaks). In many inspection applications,[9,10] only several θ slices of the straight line HT are necessary to complete the desired inspections. Real-time digital implementations of the HT are too costly and slow to allow the HT to be used in industrial inspection. There are many optical realizations of the HT which can implement the HT at high data rates with, comparatively, inexpensive architectures. We have selected the HT CGH architecture for use in our inspection processor for reasons we now note.

The HT CGH is an inexpensive and simple way to realize several slices of the HT in parallel in real-time. Its primary advantages are its simplicity (no moving parts as in other optical implementations[11,12,13]), low cost, and that the CGH is easily replicated. With our HT CGH (a film mask with a binary pattern recorded on it), multiple θ slices of the HT are formed in parallel one focal length behind the CGH.

2.1 CGH HT

The Hough transform $f_h(p,\theta)$ we consider for an input $f(x,y)$ is the (p,θ) or (normal, angle) HT

$$f_h(p,\theta)=\int\int f(x,y)\,\delta(p-x\cos\theta-y\sin\theta)\,dx\,dy, \tag{1}$$

where p is the normal distance between a line and the origin, and θ is the angle the normal makes with respect to the x-axis. In our realization, θ ranges from 0° to 180° and p is bipolar. Our CGH implementation of this HT does not use a direct mapping of input pixels to sinusoids in the HT plane as do other[14,15] CGH implementations. Instead, we encode many different orientations of pairs of orthogonal cylindrical lenses on one CGH. One lens provides integration and the second lens in each pair provides imaging in the orthogonal direction. Each encoded lens orientation realizes a different HT θ slice corresponding to the orientation θ of the lens (i.e. $f_h(p,\theta)$ at a given θ). Hence, multiple θ slices of the HT will be generated in parallel in a polar format at the output. Since the CGH is recorded with only binary data, it is also easily replicated.

Figure 1 shows the optical system used in the HT CGH configuration. The image is scanned using a red-green-blue (RGB) camera and fed to an LCTV at P_1 (the entire LCTV assembly consists of an input polarizer, or $\lambda/2$ plate, to rotate the polarization of the laser and an output analyzer). Lens L_1 Fourier transforms (FTs) the input and the spatial filter at P_2 passes only one FT order (the spatial filter also provides edge enhancement as described in Section 2.3, hence it is noted as an EE CGH). Lens L_2 inverse transforms the spatially-filtered frequency plane back into a continuous spatial image at P_3. The CGH HT is placed at P_3 and the HT slices form one focal length behind it off-axis at P_4.

Our CGH consists of different cylindrical lenses, i.e. complex transmittance functions $exp\{j\pi x^2/\lambda f_L\}$ to produce N slices of the HT at different θ_i, the CGH is the sum of N integrating cylindrical lenses rotated at different angles θ_i (with respect to the x axis),

$$\sum_{i=1}^{N} exp\{\frac{j\pi}{\lambda f_{L1}}(x\cos\theta_i+y\sin\theta_i)^2\}, \tag{2}$$

and the sum of N orthogonal imaging lenses

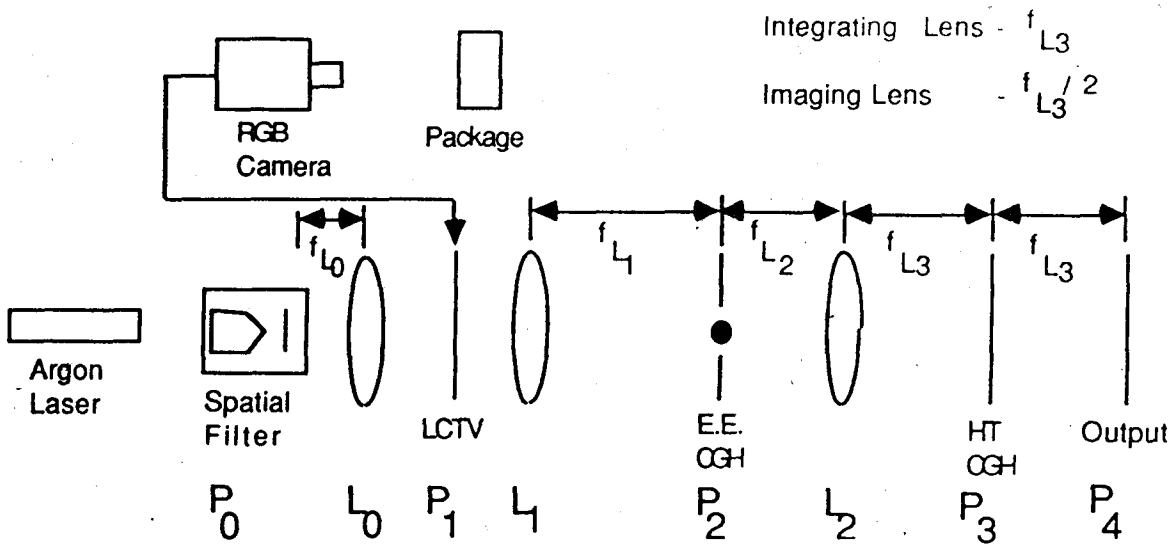

Figure 1: Real-time optical CGH HT architecture

$$\sum_{i=1}^{N} exp\{\frac{j\pi}{\lambda f_{L2}}(-x\sin\theta_i + y\cos\theta_i)^2\}, \tag{3}$$

with $f_{L2}=0.5f_{L1}$.

We now discuss the CGH encoding scheme used. We desire high accuracy with diffraction efficiency being less important (since the output is not used as the input to another optical system). We also desire a modest space bandwidth (to allow it to be easily printed) and binary encoding (to allow easy copying). As noted by Allebach[16], the most accurate representation of a function on a CGH occurs when every grid point on the CGH represents one sample point of the function (rather than allowing several grid points to represent the gray level of one sample point). Thus higher input sampling rates are preferable for more accurate representations with fewer samples. For these reasons, we choose a binary interferometric[17] CGH (in which the amplitude and phase of a carrier are modulated). To record the general complex function $a(x,y)e^{\phi(x,y)}$, the transmittance function is

$$t(x,y)=0.5+0.5a(x,y)\cos[\phi(x,y)+2\pi kx\sin\psi], \tag{4}$$

where it is assumed that $a(x,y)$ has been normalized to a maximum amplitude of 1, ψ is the angle at which the reconstruction is deflected from the axis (the last term defines the carrier spatial frequency $(\sin\psi)/\lambda$) and $k = (2\pi/\lambda)$ is the wave number.

To binarize (4), we use modulated (1-D) error diffusion[18]. This provides the most accurate CGH with the least space bandwidth product (SBWP). Error diffusion was originally suggested for improving the output of digitally half-toned pictures.[19] Similar techniques can be used when recording CGHs to attain high accuracy with less space-bandwidth requirements. The principle of error diffusion is as follows. Since we are binarizing a function, an error $\epsilon(x,y)$ is associated with each (x,y) pixel recorded on the CGH. For example, if the transmittance function $t(x,y)$ is 0.8, we threshold to 1.0 and the error $\epsilon(x,y)$ is -0.2. Instead of disregarding this error, we propagate it to the neighboring pixels and we continue this procedure for all pixels. In the 1-D modulated error diffusion technique we use, we propagate the error at each pixel to only one neighboring pixel (1-D), however, the diffused error is modulated with a spatial carrier before it is propagated. This controls the portion of the output where the error due to quantization is minimized.

Our initial[8] CGHs were printed on a 300 dots per inch laser printer and photo-reduced onto a 15 x 15 mm^2 piece of film. The CGHs described in this thesis were printed directly onto film with 2540 dots per inch (100 dots per mm) resolution with a Linotronix laser printer. The Linotronix prints 20 μm dots on 10 μm centers (oversized dots are printed since circular pixels are being used on a rectangular grid). We have developed a CGH encoding technique to compensate for dot size effects, but this was not required for our present application. The maximum size pattern that the Linotronix can print is 8.5" x 11". Significant improvements were realized because of the increased printer resolution and the elimination of the photo reduction stage.

We produced a 10 θ slice CGH on a 15x15 mm^2 film with 1536x1536 recorder spots with f_{L1}=2m and λ=488 nm. The output for a plane wave input is shown in Figure 2. Its optical efficiency was 1.4% (this can be increased by bleaching and phase recording, but it was sufficient for our purposes). The spot sizes produced were 65μm (in agreement with the theoretical diffraction limit expected). We placed a single slit (line) in the input and shifted it over the full 15mm aperture and determined that the uniformity of the output HT peak varied by as little as 14% (10% of this is due to the uniformity of the input optical beam used). This 4% variation in HT peak intensity with position of the input line is not of concern since the percentage error is known and fixed at each location (and in our inspection application, the locations of the input lines are known). The peak to sidelobe ratio (PSR) of the HT (the ratio of an HT peak value to the value between two HT slices) is 65:1 and the overall SNR (the exact expected full 10 slice HT pattern divided by the sum of its errors at all points in the HT plane) is 10. The HT CGH produced provides a 242x242 sample HT plane in 15x15 mm^2 (15mm/65μm=242). The size "a" of the CGH aperture (and the size of the HT plane with a 1:1 imaging (CGH lens), its f_{L1} and λ are related to the number of HT samples m (in a distance "a") by $a=(\lambda f_{L1} m)^{0.5}$.

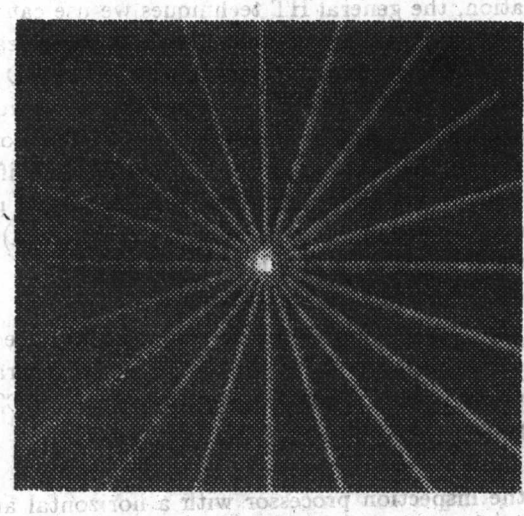

Figure 2: Output of 10 slice HT CGH when illuminated with a plane wave

2.2 LCTV

A color Epson Elf LCTV was used. It has 240 pixels horizontally and 220 pixels vertically. Only the blue channel was used (as it gave the best contrast ratio[20] for the cigarette packages) and could be read out with the blue λ=488nm Argon laser light used. With other lasers, the blue line from the camera can be fed to a different channel[20] on the LCTV and read out in another color of laser light. In one color, the LCTV gave 220x80 pixel resolution. We use the LCTV rotated 90° and hereafter refer to the 220 resolution as horizontal. The spacings s_x and s_y between liquid crystal pixels determines the maximum vertical ($1/2s_y$=1.03 cy/mm) and horizontal ($1/2s_x$=3.8 cy/mm) resolution and the spacing $\lambda f_L/s$ between replicated FT orders (0.4mm vertical

and 1.55mm horizontal). The full liquid crystal aperture (a_x=28.6mm horizontally and a_y=38.4mm vertically) determine the FT spot size (11x13 μm). The above distance calculations used f_L=795mm and λ=488mm.

2.3 Edge Enhancement Spatial Filter

The CGH at P_2 is opaque in the center, transmits light in a donut band pass region and at only certain angular orientations of FT data. It performs edge enhancement by blocking the central 50 μm at P_2. This is about 4 times the dc FT spot size. This blocks dc and low spatial frequencies below 0.13 cy/mm and thus performs edge enhancement. The CGH also contains wedge shaped slits extending $\pm 5^\circ$ on each side of each HT angle θ (this passes only input lines oriented within 5° of the desired angles θ). The CGH also passes only one FT order (the central order). This is achieved by making the P_2 CGH opaque beyond a radius of 0.4mm. This passes most of one FT order (spatial frequencies above 1.03 cy/mm are blocked, this results in blurred horizontal lines in the input), but the full 220 horizontal resolution input sampling is still attained. We also adjust the polarization of the input light, the brightness control on the LCTV, and the angle of the LCTV analyzer to achieve a remapping of the input video voltage signal to output light intensity[21]. This produces the desired white (transparent) edges on a dark background. This is not achieved by other liquid crystal phase techniques[22]. This intensity remapping further enhances our selective spatial filtering edge-enhancement.

3 Inspection Case Study

The case study used to evaluate our HT CGH inspection processor is the inspection of cigarette packages. Although this is a specific application, the general HT techniques we use can be applied to a range of inspection applications. A representative cigarette package (labelled with the key features) is shown in Figure 3. The closure stamp is labelled A, the tear tape B, and the lines labelled C and D are the diagonal lines (at 38° and 142°) which will be used to determine if the printing on the label is skewed. Specific defect inspection tasks are to determine if the tear tape is present and properly aligned, if the closure stamp is present and straight and does not extend down too far, if the label is skewed (rotated), and if there are blemishes on the label. Figure 4 shows examples of three of the defects. Figure 4a is a blemished package, Figure 4b contains a misaligned stamp, and Figure 4c has a skewed label. We only require 3 (or 4) slices of the HT (at 0°, 90°, and 38° or 142°) for this application. We used all four θ HT slices.

Production line speeds mandate that the packages be inspected at speeds of 700/minute (11.7/sec). Our real-time inspection processor is limited in speed by the LCTV which operates at TV frame rates, so we can conceptually perform 1800 inspections/minute. With the active matrix LCTV used, and the possibility of 30 ms image decay rates, we can approach 30 frames/sec rates.

The package is presented to the inspection processor with a horizontal and vertical positioning accuracy of $\pm$ 0.5mm and a rotational accuracy of $\pm 1^\circ$. The allowed skew tolerance for the package printing is 1.8° and for the closure stamp is 3.2°. We now address how the required inspection is performed. Figure 5a shows the HT of a generic package and Figure 5b shows the HT polar format. The θ=38°, 90° (horizontal lines), and 0° (vertical lines) HT slices are shown in Figures 5c to e. The 38° slice has a peak denoting the presence and location of the diagonal line D. The presence of a peak and its location denotes that the package label is present and properly aligned. The 0° slice has 4 peaks (the left and right edge of the package and the left and right edges of the stamp). The latter two smaller peaks denote that the stamp is present and aligned and in the proper location. The 90° slice has peaks (from right to left) corresponding to the top of the package, the top and bottom of the tear strip, the bottom of the stamp, and the bottom of the package. These note the presence and location of the tear strip and the stamp. On the standard Marlboro package, the banner produces a peak on the 0° slice and also provides information on the centering of the package label.