Hybrid processor to compute invariant moments for pattern recognition

David Casasent and Demetri Psaltis

Department of Electrical Engineering, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213

A hybrid optical–digital processor is described that computes invariant moments of an image in parallel by spatial convolution with a single fixed mask.

Introduction

The use of invariant moments[1,2] for pattern recognition and scene matching has received considerable attention.[1-8] The majority of this work has utilized digital processors.[3-6] In this Letter, we describe a hybrid optical–digital system that can compute all the invariant moments of a 2-D image in parallel by spatial convolution. This is an extension of earlier work[7] that required separate masks for each moment computation and differs from other optical methods[8] that use differentiation and Fourier transformation.

Invariant Moments

From the theory of invariant moments,[1-6] it is well known that the following functions are invariant under translation and rotation:

$$\phi_1 = \mu_{20} + \mu_{02}, \tag{1}$$

$$\phi_2 = (\mu_{20} - \mu_{02})^2 + 4\mu_{11}{}^2, \tag{2}$$

$$\phi_3 = (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2, \tag{3}$$

$$\phi_4 = (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2, \tag{4}$$

$$\phi_5 = (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] + (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2], \tag{5}$$

$$\phi_6 = (\mu_{20} - \mu_{02})[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] + 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03}), \tag{6}$$

$$\phi_7 = (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] - (\mu_{30} - 3\mu_{12})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2]. \tag{7}$$

In Eqs. (1)–(7), the μ_{pq} are the central moments, and they are defined as

$$\mu_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f\left(x + \frac{m_{10}}{m_{00}}, y + \frac{m_{01}}{m_{00}}\right) dx\,dy. \tag{8}$$

The μ_{pq} can be expressed [Eqs. (1)–(6)] in terms of the moments

$$m_{pq} = \iint_{-\infty}^{\infty} x^p y^q f(x,y) dx\,dy \tag{9}$$

of an image $f(x,y)$, where $p, q = 0, 1, 2, \ldots$. For instance, $\mu_{00} = m_{00}$, $\mu_{20} = m_{20} - (m_{10}{}^2/m_{00})$. A set of scale, rotation, and translation invariant moments can be obtained by substituting

$$\eta_{pq} = \mu_{pq}/\mu_{00}{}^y, \tag{10}$$

where $y = 1 + (p + q)/2$, for μ_{pq} in Eqs. (1)–(7). The proof of the invariance of these moments is given in Ref. 1. Maitra[6] has recently extended this work to include invariance to contrast differences and with attention to increased computational accuracy. From this brief introduction, we see that the set of seven absolute invariant moments ϕ_n can be used in the description or characterization of an image and that these moments can be computed from the standard moments m_{pq}.

Previous experimental work has indicated that these seven moments permit recognition of multisensor imagery[5] and provide adequate discrimination between letters.[1] To compute eleven moments m_{pq} for a 512 × 512 image in 30 msec requires at least 1.6×10^8 operations per second. This high computational load and the associated analog-to-digital converter and storage requirements make real-time operation of this system using digital technology difficult. In this Letter, we consider the use of the real-time and parallel-processing features of optical systems to implement such a pattern-recognition system.

Hybrid Processor

A simplified version of a hybrid processor to compute the ϕ_n moments from the m_{pq} moments is shown in Fig. 1. The output-wave amplitude at P_3 is given by
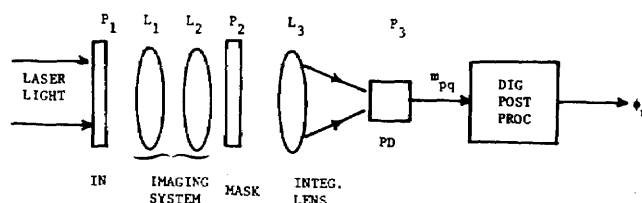

Fig. 1.   Schematic diagram of a conceptual hybrid processor to compute the invariant moments. (PD, photodetector.)

$$u(\omega_x,\omega_y) = \iint_{-\infty}^{\infty} f(x,y)g(x,y)$$
$$\times \exp[-j(\omega_x x + \omega_y y)]dxdy, \quad (11a)$$

and its value on-axis is

$$u(0,0) = \iint_{-\infty}^{\infty} f(x,y)g(x,y)dxdy, \quad (11b)$$

where $f(x,y)$ and $g(x,y)$ describe the transmittance of the input at P_1 and the mask at P_2. If $g(x,y) = 1$ or x or y or xy, the corresponding $u(0,0)$ output is m_{00}, m_{10}, m_{01}, m_{11}, etc. From this simplified initial system description, we see how the moments m_{pq} can be optically computed at P_3. A dedicated digital postprocessor can then compute the absolute invariant moments ϕ_n by appropriate addition, subtraction, multiplication, and division of the m_{pq}. From tests made on typical imagery (500 × 500 pixels in a 25-mm × 25-mm format), the dynamic range of the ϕ_n was found to be far larger than for the m_{pq} (>200 dB versus 40 dB).[7] Because of this observation and since the operations required in computing the ϕ_n from the m_{pq} can easily be performed digitally to the necessary accuracy, the hybrid-system architecture of Fig. 1 was chosen. In this system, the m_{pq} (which require the bulk of the computational load) are computed optically, and the ϕ_n are then digitally calculated from these m_{pq} values.

In the practical realization of a system of the form shown in Fig. 1, various issues arise: (1) The effective transmittance function for the mask must be bipolar (e.g., x and y are integrated over positive and negative values). Otherwise a dc bias will appear on the detector that will increase its dynamic range requirements and complicate the digital postprocessing. (2) In the system described in Fig. 1, separate masks are required for each m_{pq} computation. This greatly reduces the potential speed and parallel-processing advantages of optical processors. We now consider a more realistic version of the system of Fig. 1 without the aforementioned disadvantages. This new version does impose more-strict space–bandwidth requirements on the Fourier-transform lens. However, this is normally not the limiting factor.

If we let $g(x,y) = \exp[x \exp(j\omega_0 x)]\exp[y \exp(j\omega_0 y)]$ in Eq. (10a), the P_3 light distribution evaluation at $(\omega_x,\omega_y) = (p\omega_0, q\omega_0)$ becomes

$$u(\omega_x,\omega_y)|_{\omega_x=p\omega_0,\omega_y=q\omega_0} = u(p\omega_0,q\omega_0) = m_{pq}/p!q!$$

$$= \sum_{i=0}^{\infty}\sum_{k=0}^{\infty} \iint_{-\infty}^{\infty} \frac{x^i y^k}{i!k!} \exp[+j(i\omega_0 x + k\omega_0 y)]$$
$$\times f(x,y)\exp[-j(\omega_x x + \omega_y y)]$$
$$\times dxdy|_{\omega_x=p\omega_0,\omega_y=q\omega_0}, \quad (12)$$

where use has been made of the identity

$$\exp[x \exp(j\omega_0 x)] = \sum_{n=0}^{\infty} x^n e^{jn\omega 0x}/n!. \quad (13)$$

From Eq. (12), we see that the P_3 pattern contains the Fourier transforms of $f(x,y)x^p y^q$ for all (p,q) combinations centered at positions $(\omega_x,\omega_y) = (p\omega_0, q\omega_0)$. The value of the P_3 pattern at these locations is proportional to the desired moments m_{pq}. Each m_{pq} is weighted by a factor of $1/p!q!$. Thus at higher orders
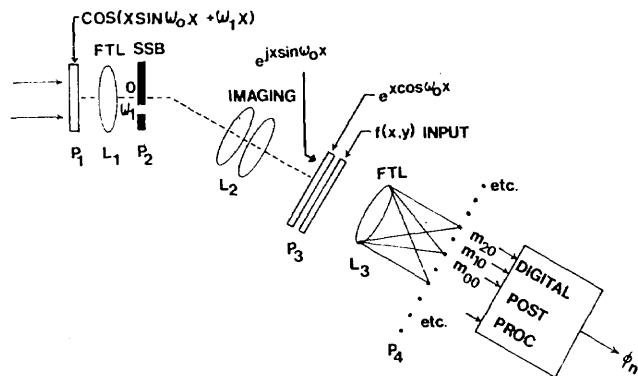

Fig. 2. Schematic diagram of a practical hybrid processor to compute the invariant moments. (FTL, Fourier-transform lens; SSB, single sideband.)

there will be less optical power or lower signal-to-noise ratio, and as a result the dynamic range of the measurement will be lower. To permit detection of all the moments with comparable dynamic ranges, the lower orders can be attenuated to avoid saturation of the detector while the higher orders remain unattenuated. In addition, relatively longer detector integration can be used at the higher orders, and the mask $g(x,y)$ can be appropriately modified to distribute the optical energy more uniformly among the orders. To avoid overlap of adjacent orders, we choose ω_0 larger than the bandwidth of the function $f(x,y)x^p y^q$. The Fourier transform of $f(x,y)x^p y^q$ is proportional to $\partial^p\partial^q F(w_x,w_y)/\partial w_x^p \partial w_y^q]$, where $F(w_x,w_y)$ is the Fourier transform of $f(x,y)$. If $f(x,y)$ is band limited by w_{max} [i.e., $F(w_x,w_y) = 0$ for $\omega_{max} > (\omega_x^2 + \omega_y^2)^{1/2}$], then $f(x,y)x^p y^q$ is also band limited by ω_0. The m_{pq} moments can thus be sensed by detectors appropriately placed in P_3.

Since the required $g(x,y)$ in Eq. (12) is complex, we realize it by two masks and the single-sideband (SSB) filter shown in Fig. 2. We will now consider only a 1-D mask for simplicity. We rewrite $g(x)$ as

$$g(x) = \exp[x \exp(j\omega_0 x)]$$
$$= \exp(x \cos \omega_0 x)\exp(jx \sin \omega_0 x) = t_A(x)t_B(x), \quad (14)$$

where t_A is real and positive and t_B is complex. To realize t_B, we record

$$1 + \cos(x \sin \omega_0 x + \omega_1 x) = 1 + \{\exp[j(x \sin \omega_0 x + \omega_1 x)] + \exp[-j(x \sin \omega_0 x + \omega_1 x)]\}/2 \quad (15)$$

at P_1 of Fig. 2. Lens L_1 forms the Fourier transform of Eq. (15) at P_2, where the SSB filter at P_2 passes only the second term in Eq. (15) and the optical axis is tilted to remove the ω_1 carrier term. Lens L_2 in Fig. 2 thus produces t_B incident upon P_3, where a second transparency t_A is placed. Leaving P_3, we thus have the desired $t_A(x)t_B(x) = g(x)$ function. At P_3, we also place the image $f(x,y)$, so that leaving P_3 we have $f \times g$. Lens L_3 in Fig. 2 forms the Fourier transform of $f \times g$, so that at P_4 we find the desired m_{pq} functions spatially separated as described by Eq. (12).

Discussion and Summary

Since the moments m_{pq} are bipolar, both the phase and the amplitude of m_{pq} must be computed. Fortunately,

from Eq. (11) we see that $u(\omega_x,\omega_y)$ for a real function $f(x,y)$ is bipolar, not complex, when evaluated at (ω_x,x_y) = $(p\omega_0,q\omega_0)$. Thus a simplified detection scheme can be used, since the output (in 1-D for simplicity) is of the form $u(\omega_x)|_{p\omega_0} = m_p/p! = (|m_p|/p!)\exp(j\pi n)$, where $n = 0$ or 1 only. Specifically, we can record the interference of $u(\omega_x,\omega_y)$ and a constant off-axis plane wave $\exp(j\omega_x x_0)$. This will produce a cos $\omega_x x_0$ fringe pattern, as in holography. When evaluated at $\omega_x = 0$, the phase of this fringe pattern will be 0 or π, depending on the polarity of m_{pq}, and the modulation of the fringe pattern will be proportional to $|m_{pq}|$. Because each m_{pq} is bipolar, we need only position two detectors in the output (separated by half the period of the interference fringe pattern). To reduce the dynamic-range requirements of the detector, we can allow the dc level of the interference pattern to saturate one detector and hence read the minimum of the fringe pattern with the second detector. Thus we require only two detectors to measure each bipolar m_{pq}. One detector will always be saturated, the output from the other detector will be proportional to $|m_{pq}|$, and which detector has the nonsaturated output determines the sign of m_{pq}.

Many other issues (beyond the scope of our initial work) deserve attention in the use of invariant moments for pattern recognition and in the optical computation of the moments. In most previous work in this area,[3–8] only seven ϕ_n moments and up to third-order moments m_{pq} have been considered. The use of additional moments will certainly enhance the discrimination of such a pattern-recognition system. Further analysis and experimentation on specific image data bases are needed to determine the number of moments needed. Another issue of concern in any optical processor is its accuracy. Since the data planes in an optical system are continuous, we expect minimal interpolation errors[6] in the hybrid processor described. However, the dy-namic-range requirements for the m_{pq} and ϕ_n moments require more extensive analysis and experimental verification. The accuracy with which the $g(x,y)$ mask must be recorded is another potential error source. However, since this mask is fixed, it should be possible to synthesize it off-line to sufficient accuracy by using a computer-controlled film recorder.

In this Letter, we have described a hybrid processor that optically computes all the bipolar moments m_{pq} of a 2-D image in parallel, from which the invariant moments ϕ_n can be computed by a dedicated digital postprocessor. The use of this and similar optical pattern-recognition techniques that do not employ matched spatial filtering appears to be of importance in the realization of practical hybrid pattern-recognition systems.

References

1. M. K. Hu, IEEE Trans. Inf. Theory IT-8, 179 (1962).
2. G. B. Curevich, *Foundations of the Theory of Algebraic Invariants* (P. Noardhoff, Groningen, The Netherlands, 1964).
3. R. C. Gonzalez and P. Wintz, *Digital Image Processing* (Addison-Wesley, Reading, Mass., 1977).
4. S. Dudani, K. Breeding, and R. McGhee, IEEE Trans. Computers C-26, 39 (1977).
5. R. Wong and E. Hall, Comput. Graphics Image Des. 8, 16 (1978).
6. S. Maitra, Proc. IEEE 67, 697 (1979).
7. D. Casasent and D. Psaltis, Proc. Soc. Photo-Opt. Instrum. Eng. 201, 107 (1978).
8. M. R. Teague, Appl. Opt. (to be published).