

# 旋转、比例、平移不变的图形识别系统

宋 建 中

**摘要:** 在高分辨率、高灵敏度的空间光调制器开始问世的今天, 降低光学相关器对图像失真的敏感程度是实时光学图形识别系统能否实用的关键。

本文叙述了一种用坐标变换的方法来实现旋转、比例、平移不变的图形识别系统。在数学上证明了傅里叶变换加上Polar-Ln变换(简称P-L变换), 可以实现平移、旋转和比例的不变性, 并进了计算机仿真。

本文提出了两种实现P-L变换的方法。一种是采用特殊的电视摄像机。用对数形的锯齿波按辐射状扫描摄像管的靶面, 另一种是采用一个计算机产生的两值光学元件, 并给出了该光学元件的相位函数的解析表达式。

本文还给出了用这两种办法实现的光学系统及实验结果。实验结果表明, 两种系统都能在输入图像发生旋转、比例变化和平移的情况下, 实时地做到对输入目标的识别。

## 一、引 言

在光学图形识别系统中(OPR), 相关技术是相当重要的。它具有较高的处理增益、平移不变而且可以同时识别多个目标。在光学上又很容易实现。一个光学相关系统如图1所示。 $P_1$ 面放输入图像,  $P_2$ 面放匹配空间滤波(MSF), 在 $P_3$ 面就可以测到相关峰值。

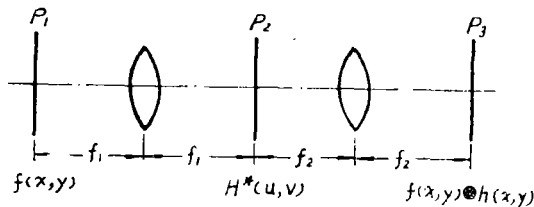


图1 光学相关系统

如果在 $P_1$ 面的输入为 $f(x, y)$ , 它的傅里叶变换为 $F(u, v)$ ; 要识别的图形是 $h(x, y)$ , 它的傅里叶变换是 $H(u, v)$ 。那么用 $H^*(u, v)$ 做匹配空间滤波器, 星号“\*”表示共轭。在图1系统中 $P_3$ 面上可以测到输出值为 $F\{F(u, v) \cdot H^*(u, v)\} = f(x, y) \otimes h(x, y)$ 。这里 $F\{\}$ 表示傅里叶变换运算, “ $\otimes$ ”表示相关运算。

但是输出的相关峰值对于输入图像和匹配空间滤波器之间的匹配程度是相当敏感的。也就是说如果 $f(x, y)$ 与 $h(x, y)$ 的差别有稍许增加, 输出的信噪比就迅速下降, 即噪声增大, 相关峰值下降。这也是相关技术的主要缺点。

克服相关技术的缺点, 又保持它图形识别的优点是研究光学图形识别系统的一个重要课题。现在已经有些方法可以消除或降低相关对失真的灵敏度。比如对图像进行某些变换后再去相关, 或采用合成判别函数来做匹配空间滤波器等等。合成判别函数也就是把输入函数的

若干失真形式进行加权，迭加成的新的函数。

本文叙述采用坐标变换的方法来实现旋转、比例、平移不变的光学相关系统。对图像做傅里叶变换后，其强度分布不随图像的平移而变化。对图像做Polar变换后，其强度分布不随图像的旋转变换，只有上、下平移。对图像做Mellin变换，其强度分布不随图像比例变化而变化，只有左、右平移。而相关本身具有平移不变性，所以如果把傅里叶变换、Polar变换和Mellin变换结合起来，用变换后的图形做相关的输入就大大降低了相关对失真的灵敏度。

傅里叶变换是透镜的一个特性，在相干光系统中，放在透镜前焦面上的图像，其傅里叶变换就在透镜的后焦面上。这里对Polar变换和Mellin变换加以说明。

Polar变换就是将原x-y直角坐标系变成新的r-θ直角坐标系，其变换关系如下：

$$r = \sqrt{x^2 + y^2}$$

$$\theta = \tan^{-1}\left(\frac{y}{x}\right)$$

在x-y坐标系的一个圆，变到r-θ坐标系就成了一个直线段。也就是说，如果在x-y平面

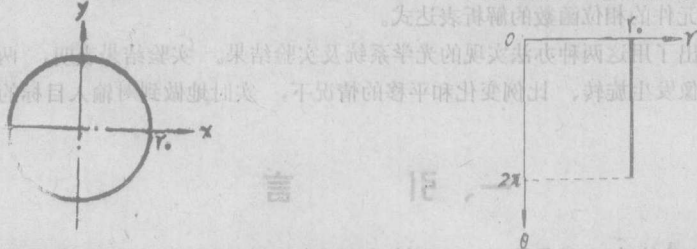
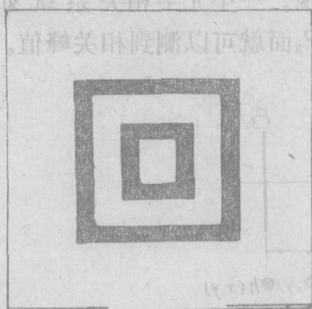


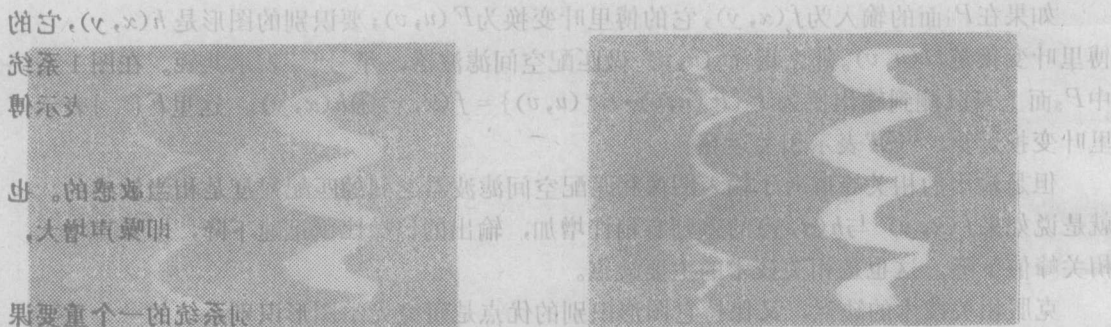
图2 Polar变换



3-1



3-3



再经变换某些失真形式... 图像旋转后，其Polar变换只有上、下平移

内, 有一点沿圆周运动, 那么经过Polar变换, 这个点在 $r-\theta$ 平面内是沿直线移动。所以, 把一个图像旋转, 它的Polar变换只发生上、下平移, 如图2到图3所示。

函数 $f(x, y)$ 的Mellin变换定义如下:

$$M(u, v) = \iint_0^{\infty} f(x, y) x^{-j^{u-1}} y^{-j^{v-1}} dx dy \quad (2)$$

令 $\xi = \ln x, \eta = \ln y$ , 并用 $F\{ \}$ 表示傅里叶变换

则(2)式可写成:

$$M(u, v) = \iint_{-\infty}^{\infty} f(\xi, \eta) e^{-j^{u\xi + v\eta}} d\xi d\eta = F\{f(\xi, \eta)\}$$

函数的Mellin变换, 就是先对函数坐标取对数后再做傅里叶变换。Mellin变换具有比例不变性, 可以证明如下: 假设函数 $f(x, y)$ 按比例变化为 $f(ax, ay)$ , 其Mellin变换为:

$$\begin{aligned} Ma(u, v) &= \iint_0^{\infty} f(ax, ay) (ax)^{-j^{u-1}} (ay)^{-j^{v-1}} d(ax) d(ay) \\ &= a^{j^{(u+v)}} \cdot k \cdot \iint_{-\infty}^{\infty} f(\xi, \eta) \xi^{-j^{u-1}} \eta^{-j^{v-1}} d\xi d\eta \\ &= k \cdot e^{-j^{(u+v)} \ln a} \cdot M(u, v) \end{aligned}$$

其中,  $k$ 是常数,  $\xi = \ln x, \eta = \ln y$

$$\therefore |Ma(u, v)| = k \cdot |M(u, v)|$$

即, 函数按比例发生变化以后, 其Mellin变换的模保持不变。

如果只取图像的强度分布, 按图4顺序做傅里叶变换、Polar变换和Mellin变换后再相关, 就得到了平移、旋转、比例不变的相关系统。因为做相关时, 是取输入的傅里叶变换与匹配空间滤波器相乘, 所以Mellin变换中的傅里叶变换可以与相关运算的傅里叶变换合二为一。所以实际上只要做图像的傅里叶变换和Polar-Ln变换。

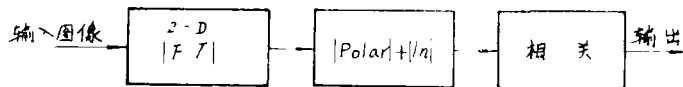


图4 平移、旋转、比例不变的相关系统方框图

## 二、计算机仿真

计算机的输入图像共有三个。图像1是一个边长为15mm的正方形。图像2是一个内径为15mm, 外径是18mm的圆环。图像3是失真的图像1, 即边长为7.5mm的正方形再旋转45°。三个图像的位子都互有平移。

把它们分别按照图4所示流程进行运算, 仿真结果列于表1中。具体仿真过程示于图6~图9。从仿真结果看出, 虽然, 图像3比图像1旋转了45°, 大小变化了1倍, 并且互有平移, 但是两者经过变换后的相关峰值仍达到了自相关峰值的86%。如果不做上述变换, 直接把图像1与图像3相关, 其相关峰值只有0.2, 也就是说只有自相关峰值的22%。可见经过坐标变换后, 相关峰值对失真的灵敏度是大大下降了。图像1与图像2, 虽经变换, 其相关峰

值仍很小, 说明不同的目标仍不相关。即, 上述变换仍可保留相关识别的特性。

表 1

	相关峰值	与自相关比	备 注
图像 1 自相关	0.908	100%	经过变换
图像 1 与图像 3	0.78	86%	"
图像 1 与图像 2	$2.05 \times 10^{-8}$	0.2%	"
图像 1 与图像 3	0.20	22%	不经变换

图像 1: ■ 图像 2: ○ 图像 3: ◆

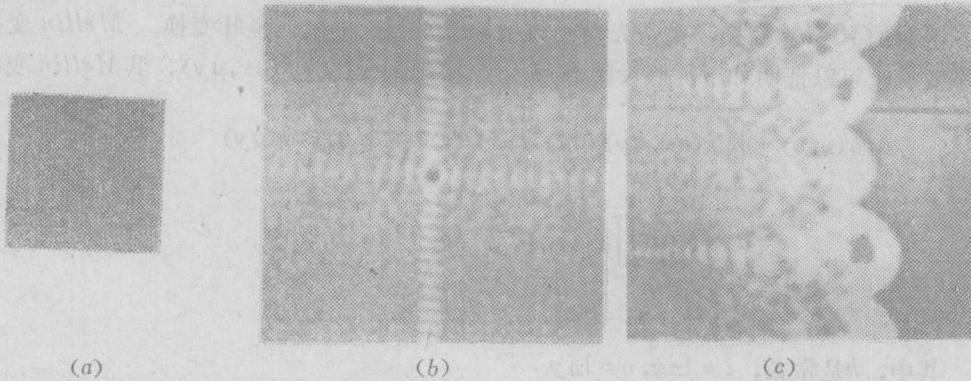


图 6 (a)输入图形, (b)输入图形的傅里叶变换 (c)输入图形经傅里叶变换后再做Polar-Ln变换

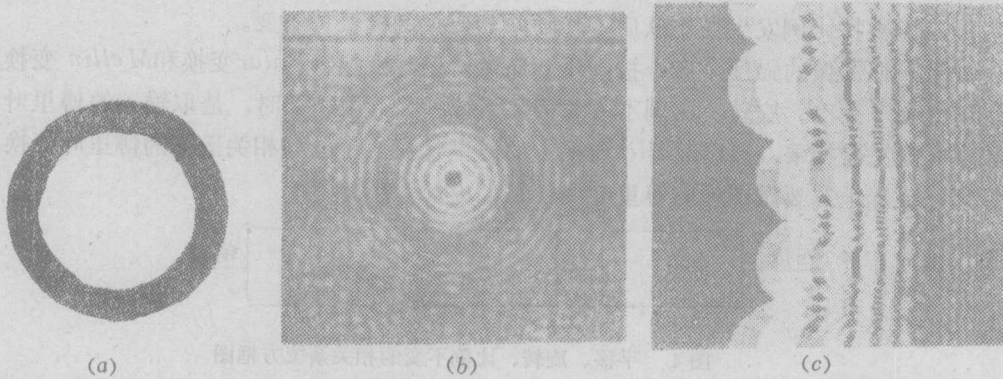


图 7 (a)输入图像(b)输入图像的傅里叶变换 (c)输入图像经傅里叶变换后再做Polar-Ln变换

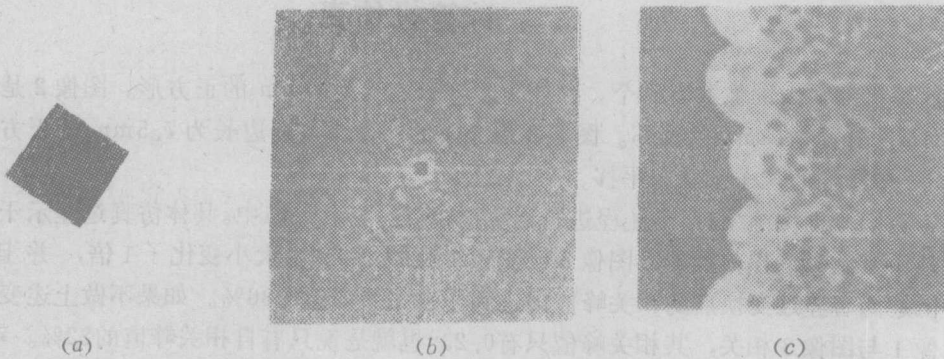


图 8 (a)输入图像(b)输入图像的傅里叶变换 (c)输入图像经傅里叶变换后再做Polar-Ln变换

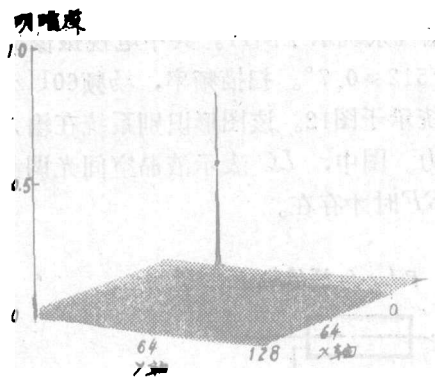


图9 图像1与图像3经变换后的相关峰

### 三、光学上的实现

#### 1. Polar-Ln电视摄像机

因为电视摄像机是一个图像传感器，所以在具体实现图像变换时，人们首先想到了用电视摄像机。早在1977年D. Casasent和M. Kraus曾设计了一个能做Polar变换的电视摄像机。该机采用像车轮的幅条那样由中心向外呈幅射状的扫描方式扫描摄像管的靶面，它的输出信号送到普通光栅式扫描的监视器。这样，这套特殊的闭路电视系就可以实现Polar变换了。

为了做Polar-Ln变换，需要在原有基础上加入Ln变换。在电路上就是把原来的线性扫描锯齿波改为对数形的扫描锯齿波。扫描电压按对数变化，输出图像的位子就按对数分布了，这可以从图10中看到。

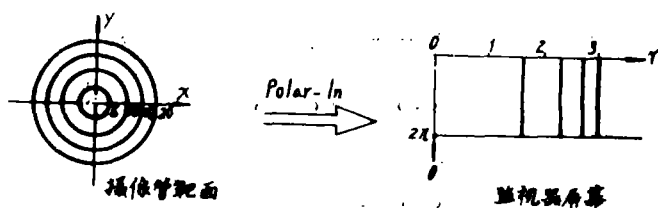


图10 用电视系统实现Polar-Ln变换

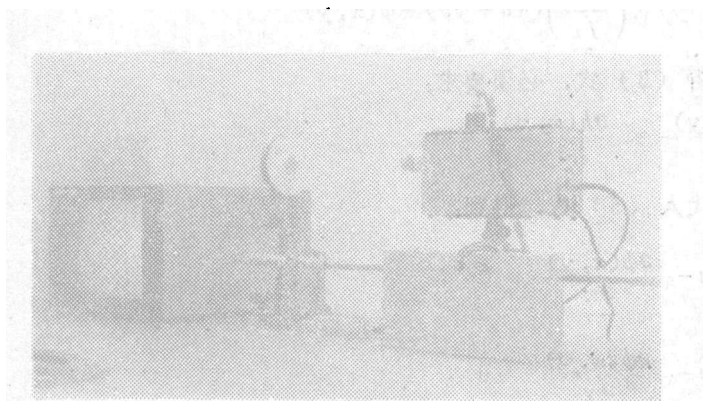


图11 Polar-Ln变换的电视系统

整个  $Polar-Ln$  变换电视系统示于图11。其中电视摄像机扫描一帧为 512 条线，隔行扫描。最大分辨角度为  $320^\circ/512 = 0.7^\circ$ 。扫描频率，场频 60Hz，行频 15.36kHz。用这套系统构成的实时图形识别系统示于图12。该图形识别系统在输入图形发生旋转、比例变化或平移时仍具有很强的识别能力。图中， $LC$  表示液晶空间光调制器。 $MSF$  表示匹配空间滤波器。参考光只有在制做  $MSF$  时才存在。

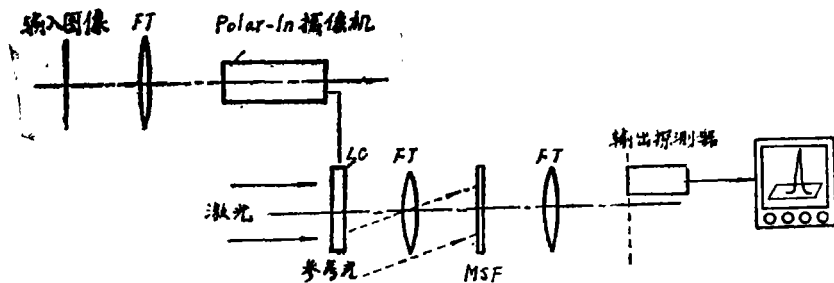


图12 使用  $Polar-Ln$  摄像机构成的实时图形识别系统

### 2. 计算机产生的光学元件 (HOE)

对图像做  $Polar-Ln$  变换，也就是对图像所在的平面进行  $Polar-Ln$  坐标变换。图13给出一种坐标变换系统。一个具有相位函数为  $\phi(x, y)$  的  $HOE$  贴在图像输入平面后面。平行光入射到输入平面，经过输入面和  $HOE$  面后再经过一个透镜的傅里叶变换，在透镜的后焦面上就是变换后的输出。

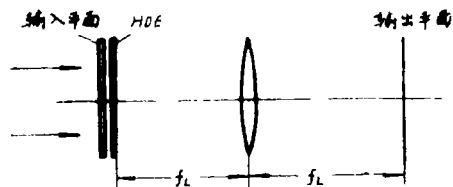


图13 使用  $HOE$  的坐标变换系统

假设输入图像为  $A_i(x, y)$ ，输出为  $A_o(u, v)$ ，则根据图13所示系统，应有：

$$\begin{aligned}
 A_o(u, v) &= \iint_{-\infty}^{\infty} A_i(x, y) e^{i\phi(x, y)} e^{i\left(\frac{k}{f_L}\right)(xu+yv)} dx dy \\
 &= \iint_{-\infty}^{\infty} A_i(x, y) e^{ih(x, y)} dx dy
 \end{aligned} \tag{3}$$

$$\text{这里, } h(x, y) = \left(\frac{k}{f_L}\right)(xu + yv) - \phi(x, y) \tag{4}$$

在数学上解析 (3) 式，必须要求：

$$\frac{\partial h(x, y)}{\partial x} = \frac{\partial h(x, y)}{\partial y} = 0 \tag{5}$$

将 (4) 式代入 (5) 式，得到：

$$\left(\frac{k}{f_L}\right)u - \frac{\partial \phi(x, y)}{\partial x} = 0 \tag{6}$$

$$\left(\frac{k}{f_L}\right)v - \frac{\partial \phi(x, y)}{\partial y} = 0$$

(6) 式可以改写为：

$$u = \left( \frac{f_L}{k} \right) \frac{\partial \phi(x, y)}{\partial x}$$

$$v = \left( \frac{f_L}{k} \right) \frac{\partial \phi(x, y)}{\partial y}$$
(7)

(7)式给出了输入和输出的对应关系, 在输入面上一点 $(x, y)$ 是按照(7)式给出的关系对应到输出面上一点 $(u, v)$ 。

现在, 要求:  $u = \ln \sqrt{x^2 + y^2}$

$$v = -\tan^{-1} \left( \frac{y}{x} \right)$$
(8)

将(8)式代入(7)式, 可得到:

$$\frac{\partial \phi(x, y)}{\partial x} = \left( \frac{k}{f_L} \right) \ln \sqrt{x^2 + y^2}$$

$$\frac{\partial \phi(x, y)}{\partial y} = - \left( \frac{k}{f_L} \right) \tan^{-1} \left( \frac{y}{x} \right)$$
(9)

(9) 式的完整解是:

$$\phi(x, y) = \left( \frac{k}{f_L} \right) \left[ x \ln \sqrt{x^2 + y^2} - x - y \tan^{-1} \left( \frac{y}{x} \right) \right]$$
(10)

(10) 式就是做 *Polar-Ln* 变换的 *HOE* 应具有相位函数。把  $\phi(x, y)$  进行编码, 制做 *HOE* 的方法很多, 如多面 *HOE*<sup>[1]</sup>, 迂回相位 *HOE*<sup>[2]</sup>, 连续相位 *HOE*<sup>[3]</sup>, 等等。使用那种方法, 决定于所要求的精度和具有的设备条件。现在  $\phi(x, y)$  的解析形式已经找到, 所以采用连续相位法最好。

用普通的全息方法记录一个相位  $\phi(x, y)$ , 是把它与一个具有  $\alpha$  角度的参考光相干涉, 记录干涉条纹。条纹的振幅分布为:

$$t(x, y) = 0.5 [1 + \cos(2\pi\alpha x - \phi(x, y))] \quad (11)$$

$t(x, y)$  的最大值位于:

$$2\pi\alpha x - \phi(x, y) = 2n\pi, \quad n = 0, 1, 2, \dots \quad (12)$$

将(10)式代入(12)式, 得到:

$$\alpha x - \frac{1}{f_L \lambda} \left[ x \ln \sqrt{x^2 + y^2} - x - y \tan^{-1} \left( \frac{y}{x} \right) \right] - n = 0 \quad (13)$$

改变  $\alpha$  角度, 可以改变所记录的条纹密度。 $\alpha$  又叫全息图的载波频率。选择  $\alpha$  值后, 用数值解法求出满足(13)式的点  $(x, y)$ , 并将同一个  $n$  值下的点逐个相连就画出了所要求的 *HOE*。

$\alpha$  的值要足够大, 才能避免不同衍射级别的图形重叠在一起。可以证明, 如果图像的最高空间频率为  $u_m$ , 要分开第一级和第二级的衍射图形, 载波频率  $\alpha$  应大于  $3u_m$ 。即,

$$\alpha > 3u_m \quad (14)$$

图14给出了计算机画出的 *Polar-Ln* 变换的 *HOE* 图形。

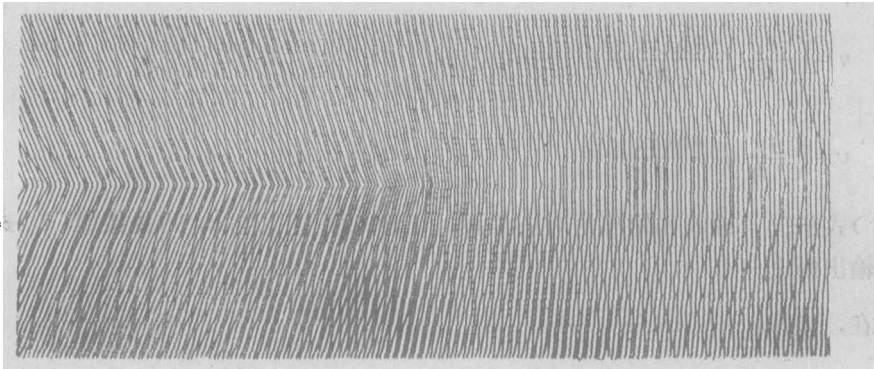


图14 计算机画的Polar-Ln变换HOE  $\alpha = 40$   $N = 200$  (用Imagen-300激光打印机画的)

用HOE构成的实时图形识别系统示于图15。在这个系统中电视摄像机是普通光栅式扫描的或像CCD摄像机那样自扫描。激光源是用激光二极管。参考光也只是在制做匹配空间滤波器时才存在。

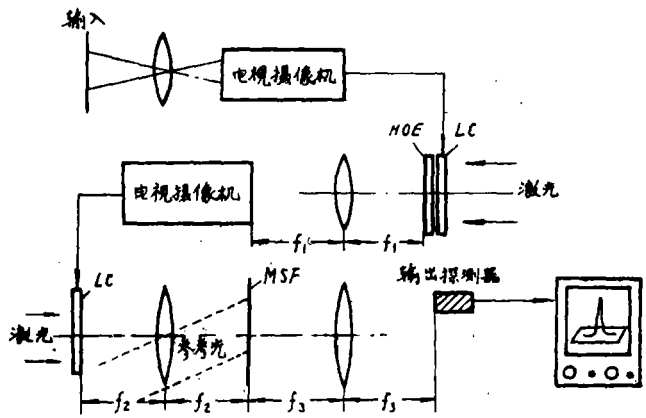


图15 用HOE构成的实时图形识别系统

### 四、实验结果

由于现有液晶光调器的分辨率太低，对傅里叶变换图形中的较高成份不能响应，所以在实验中只做了旋转和比例不变的相关，但这并不应该影响对平移不变性的承认。具体实验装置如图12和图15所示。只是输入图像在做Polar-Ln变换之前，不做傅里叶变换。

实验的图像是三个字母“X”、“T”与“O”。当把“X”在尺寸上按比例放大1.3倍，或将



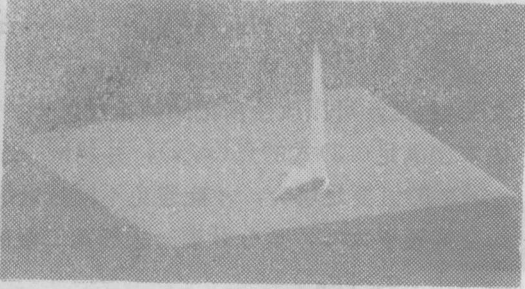
16-1

图像1：字母“X”



16-2

字母“X”的Polar-Ln变换



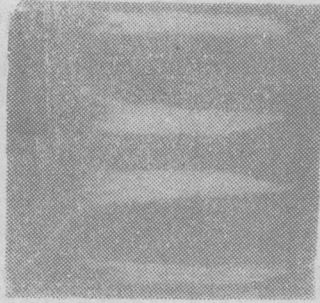
16-3

字母“X”的自相关峰



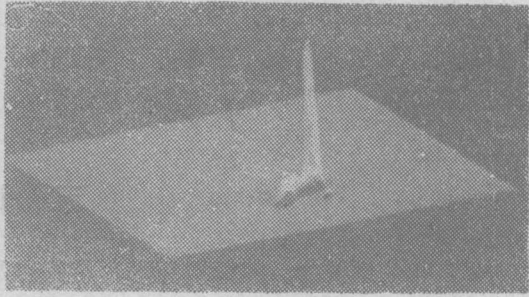
16-4

图像2: 字母“X”放大1.3倍



16-5

字母“X”放大1.3倍后的Polar-Ln变换



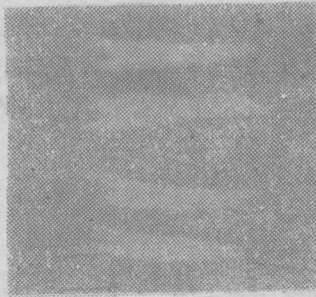
16-6

放大的“X”与原来的“X”的相关峰



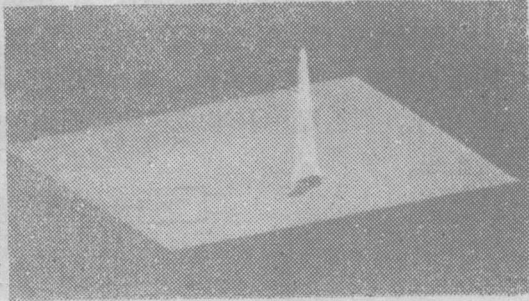
16-7

图像3: 字母“X”旋转45°



16-8

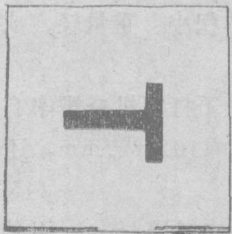
旋转的“X”的Polar-Ln变换



16-9

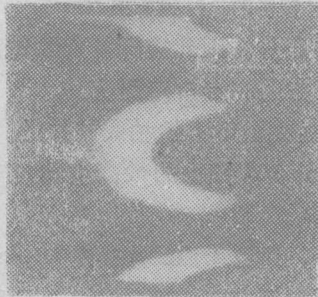
旋转的“X”与原来的“X”的相关峰

图16 旋转、比例不变的相关实验结果(1)



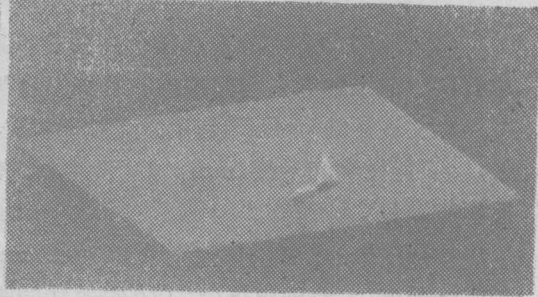
17-1

图像4: 字母“T”旋转90°



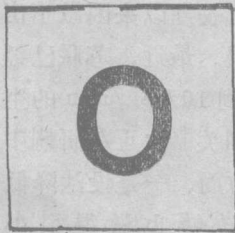
17-2

字母“T”旋转90°后的Polar-Ln变换



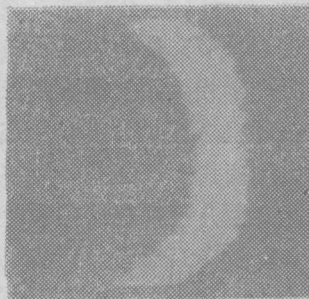
17-3

“T”与“X”的相关峰



17-4

图像5: 字母“O”



17-5

字母“O”的Polar-Ln变换



17-6

字母“O”与“X”的相关峰

图17 旋转、比例不变的相关实验结果(2)

“X”旋转大约45°时，所输出的相关峰值几乎与自相关时差不多。但当“X”与“T”或“O”去相关时，仍然只有很小的峰值。与计算机仿真的结果一致。证明这两个系统都能在图像发生旋转和比例变化时，实时识别特定的目标。实验结果示于图16和图17。

## 五、讨 论

在相关系统中、匹配空间滤波器是事先根据给定的目标做好的。在使用中只要实时输入图像，就可以进行实时相关、识别给定的目标。

采用 *Polar-Ln* 电视摄像机方法，可以得到较高的变换精度，实时相关系统中只用一块液晶。但它需要特殊的电视摄像机。体积不易做小，适用于实验室研究。

而计算机产生的光学元件方法可以使用 *CCD* 等固体传感器，使整个相关系统做得很小，适用于军事、遥感等场合。进一步改进 *HOE*，把 *Polar-Ln* 变换与傅里叶变换做在一个 *HOE* 上，那么用 *HOE* 构成的实时图形识别系统就更加有吸引力。图18给出改进后的实时相关系统。

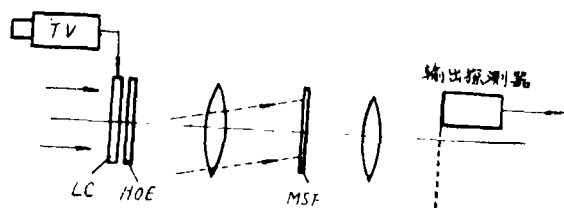


图18 改进的 *HOE* 构成的实时图形识别系统

*HOE*和*MSF*实际上都是个高分辨率的底片。因此实际装置可做到体积小、重量轻、成本低廉。在未制导、空间遥感、目标识别等方面将会有广泛应用。

本文叙述的 *HOE* 是用激光打印机先画成图后再缩拍成胶片的。由于打印机分辨率有限， $\alpha$ 只能做到40对线/mm，按照 $\alpha = 4u_\lambda$ 来计算，图像的最高空间频率只有10对线/mm，还远不能满足实际需要。

如果图像的最高空间频率是40对线/mm，那么*HOE*的载波频率 $\alpha = 160$ 对线/mm。用激光或电子束1:1地画出这个*HOE*它的线间距离只有 $1/160 = 6\mu\text{m}$ 。而刻画的精度应能达到光学表面的加工精度，因为线的位子代表了附加的相位。若以1~2波长的精度来要求，刻画*HOE*的精度应在 $1\mu\text{m}$ 左右。这样的精度，用计算机控制电子束直接对底片曝光已可以达到。

在实时相关系统中，二维的实时光调制器(*SLM*)是个关键元件。长期以来因做不出高灵敏度、高分辨率的实时空间光调制器，实时光学相关系统得不到进展。最近，苏联已经在这方面有了突破<sup>[4]</sup>。今年，日本的浜松电视公司也做出了分辨率可达到10对线/mm的空间光调制器，还可以控制输出图像进行旋转和比例变化。一些实时光学相关器也正在研制中，有的已做成了样机。这些实时光学相关器的注意力几乎都是放在两个方面，一是设法降低相关对失真的灵敏度，二是在体积上压缩，大多以激光二极管为光源、透镜也做得很小。*ERIM*制造一个4路实时光学相关器，用一个2维的*CID*做输出探测器。它的体积为 $15 \times 23 \times 42\text{cm}^3$ 。重量为8kg，功耗55W。*GDP*制造的实时光学相关器采用了计算机产生的合成判别函数做空间匹配滤波器。整个系统直径小于27cm，大约30cm长。它是放在导弹上

做实时图形识别用的。

只要空间光调制器的研制有了突破,实时相关器一定会得到广泛应用。本文所述的方法也会得到实践的证实和进一步改进。

## 六、结 论

本文叙述了用傅里叶变换加Polar-Ln变换的方法实现旋转、比例、平移不变的实时光学图形识别系统。首先用计算机仿真证明上述原理正确。然后提出了两种实现上述变换的实时系统。一种是用对数形锯齿波呈辐射状扫描电视摄像管的靶面;另一种是用计算机产生一个两值的光学元件,其相位函数表达式为:

$$\phi(x, y) = \left( \frac{x}{f_L} \right) \left[ x \ln \sqrt{x^2 + y^2} - x - y \cdot \tan^{-1} \left( \frac{y}{x} \right) \right]$$

实时光学相关器的主要障碍是高灵敏度、高分辨率的空间光调制器。现在苏联已经在这方面有所突破。实用的实时光学相关器正在研制中,很快将得到试用。

本文所叙述的工作是作者在美国Carnegie-Mellon大学工作期间在D. Casasent教授指导下进行的。作者对D. Casasent教授表示感谢。

### 参 考 文 献

- [1] H. Bartelt and S. Case, "Coordinate Transformations Via Multifacet Holographic Optical Elements" Opt. Eng., 1983, 22, 497
- [2] W. H. Lee, "Computer Generated Holograms, Techniques and Applications" in Progress in Optics, E. Wolf, Ed. (North-Holland, New York, 1978), pp. 119—232
- [3] D. Casasent and C. Szczutkowski, "Optical Mellin Transforms using Computer Generated Holograms," Opt. Commun., 1976, 19, 217
- [4] D. Casasent, F. Caimi and M. Petrov, Applied Optics, 1982, 21, Nov., 3846
- [5] D. Casasent and J. Song, "A Computer Generated Hologram for Diffraction Pattern Sampling", Proc. Soc. Photo-Opt. Instrum. Eng., 1985, Jan., 523
- [6] J. W. Goodman, Introduction to Fourier Optics (McGraw-Hill, New York, 1968)
- [7] J. Cederquist and A. Tai, "Computer Generated Holograms for Geometric Transformations," Appl. Opt., 1984, 23, 3099

## Rotation, Scale and Shift Invariant Pattern Recognition System

Song Jianzhong

### Abstract

This paper introduces the Polar-Mellin transformation for distortion-invariant pattern recognition with emphasis on the implementation of the Polar-Mellin transform by a computer-generated optical element and by a special TV camera.

When the computer-generated optical element or the TV camera for the polar-In transform is used for correlation, the pattern which is interested can always be recognized even though the pattern is rotated or scale-changed or shifted.

In this paper, two kinds of distortion-invariant systems for pattern recognition using the computer-generated optical element and the TV camera are given. Computer simulations and primary experiments have proven that methods described in this paper are correct.