

双眼立体视觉的信息加工

中国科学院生物物理研究所五室 双眼视觉研究组*

一、随机点的立体图对

动物和人生活在三维空间中,由于生存上的需要——觅食和躲避敌害——必须具有深度知觉,这也是在生活环境中长期适应的结果。

但是动物和人又是通过什么去感觉出三维空间的景物的呢?我们知道眼睛的网膜是可以感受光强的,但是眼睛网膜所感受到的光强是外界景物的三维投影,至于深度这方面的信息(如果是振动形式的信息,那么“深度”信息也就由相位差来表示。)就很难设想从单个眼睛网膜的光强信息上传送入脑。当然,一个眼睛依靠水晶体的调焦,可以由调焦的程度上知道景物的远近;也可以从经验知道东西的大小和网膜上成象大小的比例来推断出距离;也可以从三维物体本身的投影角度等等来推断。可是直接实时地由单眼网膜上就传入深度信息还没有实践上的依据和科学上的推断。一般都认为双眼能给出深度信息,的确在这方面有可靠的实践经验。例如我们观看立体投影,我们所戴上的偏光眼镜,它所起的作用就是使我们左眼只见到放映机放出的从左边拍摄的电影片,右眼只见到放映机放出的从右边拍摄的电影片。因此我们所观看的,虽然是映在平面银幕上的电影,却如同我们生活在三维环境中一样,用两个眼睛从两个角度观看立体的景物。另外,照相机上的自动对距离也是同样的原理。那么两个眼睛观察和一个眼睛观察究竟又有些什么不同呢?两个眼睛观察为什么就能将深度看出来呢?究竟“深度”这个信息是通过什么信号刺激在两眼网膜上才能使人感觉出深度的呢?

现在已经知道“视差”就是“深度”给予两眼

的信号。

首先,由于物体远近不同,在左右两眼上的投影也不同(见图1)。假如F为凝视点,这点对应于两眼的中央凹(0),则 P_1 、 P_2 、 P_3 点在左眼的投影位置和在右眼的投影的位置都不是相应点,可以说这些投影点各自在左右两眼位置上的差别就是这一点的视差。

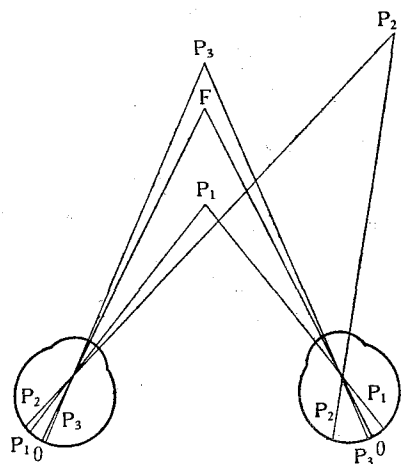


图1 双眼上的视差

视差是客观的物理现象。用有一定高度的小方块站立在一正方形纸片的正中,然后用照相机来代替眼睛,在左右两边上方各照一张相。这样左边一张底片就代表在左眼网膜上的投影,右边一张底片就是代表右眼网膜上的投影。可以见到正中小方块在两张底片上都向旁边移动了一点(见图2)。如果正中小方块愈高则两边底片上这小方块向旁边偏移得愈多。这实际上就是不同高度在左右两边所投影的视差。因此视差也是一种客观的物理现象,根据这个现

* 物理所三室四组的同志们以及计算所三室的同志曾对本工作给予大力支持和协助。

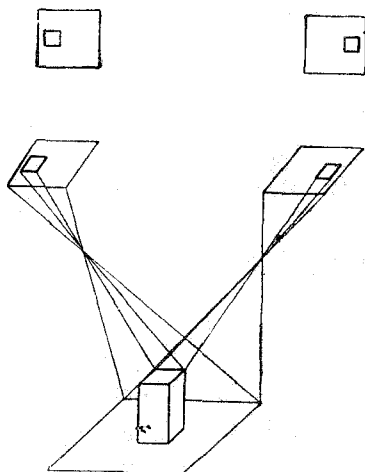


图2 视差的图示

象我们完全有根据认为客观的三维景物在两眼网膜上的投影是具有视差的。

但是我们又如何来证明有了视差就有了深度感知？视差就携带了深度信息呢？

请看封3图3，如果用一副红绿色滤片的眼镜*，那么将红色滤片放在左边，绿色滤片放在右边，就可以见到一个凹下去的椭圆抛物面，它具有一个真实的深度感觉。

这个图就是我们根据深度知觉和视差的关系，将一个三维的图型分解成左右两张二维的投影图而成的。当你左眼只看左边的那一张（现在是红色的）右眼只看右边的那一张（现在是绿色的）那么“做在”左右两张二维图中的视差就会把深度信息携带给你。

具体的做法是选择一个三维图形：椭圆抛物面 $x^2 + y^2 = 100^2(z + 10)$ ，（ z 是深度的坐标轴，在这里 $-10 \leq z \leq 15$ ）将它上面的每一点根据视差，分别的投影到左右两张图上。也就是说，每有一点 (x, y) 就有一个相应的 z 值，这时根据 z 的大小反过来对所在的这一点 z 值作为一个相应的变化，也就是在水平轴上位移一下这个点子。就是这样在数字电子计算机上将左右两眼网膜上的投影图全部都计算出来。为了避免明显的图形对成象的“先验”影响，采取了随机点质地，每 10×10 个小点就合成一个或是黑或是白的随机出现的大点，在根据视差

的大小分别投影到左右两边的图上时，这个随机点就分别的到左右两张图的相应点上。这样，单眼去观察就看不出任何图型。

由于这两张图的产生仅仅只是根据视差，而当分别用左右眼去观察这左右两张图时能融合成一个预先定好的三维图象，这说明了视差的确携带了深度信息，它在双眼体视上所起的作用是主要的。

利用这种立体图对的方法来研究深度视觉问题，最初是由匈牙利人 Bela Julesz 在美国贝尔电话实验室进行的。他的工作^[1]开始时是利用电子计算机产生两张相同的随机点的图，并将其中一张中的一些图型的随机点水平位移了一个距离，也就是造成了一对具有视差的图对。这样当用双眼去观察，就可以看到融合成的这一定的立体图型浮凸在背景上面或凹到背景下面。以后又用图型上的变化说明了：其融合的条件并不是除了视差之外两张图不允许有任何其它少许的差别，例如两眼的图的大小如果其中之一缩小一些（10%）；或是其中之一具有噪声（20%），或是一图因出焦而模糊等等，都不会影响融合。后来又用一些三维图型如螺旋线、马鞍形等等根据视差分别解析到左右两张图上去，但可以用双眼融合成一个三维图象。这种研究方法在有了显示装置的电子计算机的今天是非常方便而又卓著成效的。在视差和双眼体视的关系上所得的结果也是非常有意义的。

二、立体图对的空间频谱

上面已经说明了视差和深度的关系。客观物体的深度在双眼上必定产生视差，而有视差的两张二维图型也可以由人眼融合出一个三维的图型。可以说在双眼深度感觉中视差的确携带了深度的信息；也可以说视差是双眼视网膜上的一个深度刺激。接下去的问题是视觉系统又如何来加工这个刺激。有人在猫^[2]和猴子^[3]

* 请读者同志自己设法做一副红绿色的滤色眼镜。只要用一张红色滤片（或用红色玻璃纸——包糖的也可以）和一张绿色滤片（或用绿色玻璃纸——包糖的也可以）贴在已做成眼镜框那样的硬纸上，代替眼镜片就可以了。

视皮层上记录出了一些细胞专门对视差敏感,而且不同的细胞所对应的视差大小不同。那么深度辨认是否脑子直接将两眼网膜上的空间位置逐点比较定出视差呢?还是有其它别的加工方法呢?

最近光学信息加工方面利用空间频谱的概念^[4]。这和无线电中随时间变化的频谱概念相同。将空间的光亮强度的变化也用频谱来表示。用这种概念去看待物与象就不是按习惯的观点把它们看成是一群明暗点的组合,而是把它们看成为一群明暗波状结构的组合。这种波是空间的函数,具有空间频率。例如,每两条线间隔为 0.01 毫米的光栅,其空间频率为每毫米 100 周。这种概念也波及到生物学的领域,尤其在研究视觉信息加工方面,有人^[5]从实验结果中推论视觉神经系统是由多道通频带组成,有人^[6]从实验证明大脑就是一个频率分析器。也有人^[7]从猫视网膜神经节细胞的实验中解释感受野具有对不同空间频率敏感的特点等等。这些利用空间频谱的分析来研究视觉信息加工的过程,并把视觉信息加工过程看做是对图象频谱加工的观点已越来越吸引人们的注意了。

就拿双眼深度视觉实验来说,过去用视差作为刺激,而现在有不少人用条纹来刺激并且发现,双眼观察不同频率的条纹会引起体视感。象左眼观察黑白相间的条纹 9 条/度(张角角度),右眼观察 8 条/度,则可见到一排黑白的栏杆直立斜向前面伸展出去。左边靠人近右边离人远^[8]。当用此法刺激,同时两眼都能看到一旋转的圆柱,结果可以看到一个真实的黑白相间的条纹组成的圆柱,围绕了一个垂直轴在转,并且总是向观察低频条纹的那一眼的方向转动^[9]。因此虽然视差携带了深度信息这是目前不少人所公认的,而两眼观察不同空间频率的条纹也会引起深度感觉,这也是一个事实。

假如说视觉信息的加工是根据分析图象的空间频谱来着手,那么在深度视觉中视差和空间频谱之间又有什么关系?深度是否是直接由脑子将左右两眼逐点的视差计算出来而感知的?还是通过“频谱分析”将左右两眼的空间频

谱比较而得到的呢?究竟视差和空间频谱之间是否存在一定的关系?针对这个问题我们进行了一些工作。

我们的工作是将一对有视差的立体图对,利用富利埃变换计算出它们的频谱,然后再比较左右两眼立体图对在频谱上所表现出来的差别。

我们所用的具有视差的立体图对是上面报道中所叙述的那个椭圆抛物面

$$x^2 + y^2 = 100^2(z + 10)$$

(其中 z 是深度轴)和 $y = 0$ 相截的那一条线上 x 自 -500 到 0 那一部分的立体图对(见图 4a 左, 4a 右),经计算机做富利埃变换得到其频谱(见图 4b 左, 4b 右)。

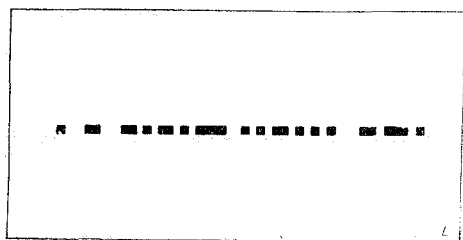
由于随机点的图型对于进一步的分析增加了复杂性;多了一个难于控制的因素,所以我们又进行了有规则质地的图型,其两种状态的分布采取每连续 9 个点 $f(x) = 0$, 然后接着 1 个点 $f(x) = 1$, 这样的图型质地见图 5a。这种图型质地的抛物线

$$\begin{cases} x^2 + y^2 = 100^2(z + 10) \\ y = 0 \end{cases}$$

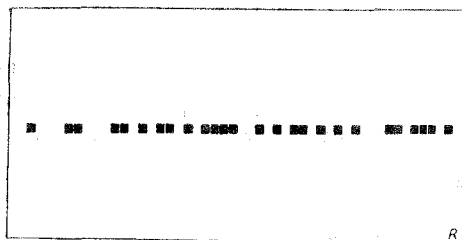
当 x 自 -500 到 0 时,再根据 z 的大小换算为相对应的视差就可得到左右两眼的投影(见图 5a 左, 5a 右),在电子计算机上利用快速富利埃变换,就得到了它们相对应的频谱(见图 5b 左, 5b 右)。

从所得的空间频谱来看,最有意义的是左眼投影的频谱 5b 左其高峰比没有深度均匀图型的频谱 5b 的高峰向高频处位移了一些,而右眼投影的频谱 5b 右其高峰比没有深度均匀图型的频谱 5b 的高峰向低频处位移了一些。这些偏移现象在随机点的图型的频谱上 4b 左、4b 右也是相同的。

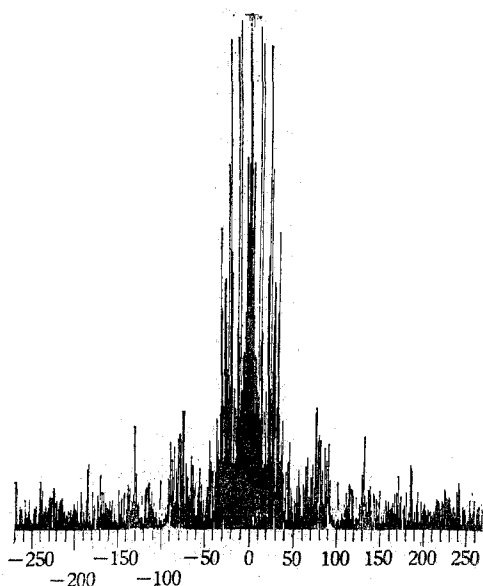
如果你用红片在左,绿片在右去观察封 3 图 6 或图 7,则你可以见到半条抛物线左边离你近,右边离你远,而这些图型的频谱又正巧是左边的高峰向高频移,右边的高峰向低频移。这和有人做的心理物理实验的结果相同。当左眼看高频条纹,右眼看低频条纹,其融合图型是从



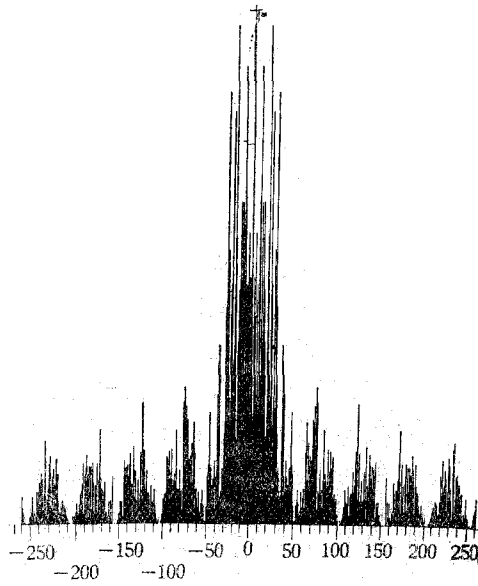
4a 左



4a 右



4b 左



4b 右

图4 随机点质地的抛物线

$$\begin{cases} x^2 + y^2 = 100^2(z + 10) \\ y = 0 \end{cases}$$

$-500 \leq x < 0$ 在左右两眼的投影及其空间频谱

上左: 左眼的投影
下左: 上图的频谱

上右: 右眼的投影
下右: 上图的频谱

左边逐渐向右前方伸展出去的一排黑白相间的“栏杆”。如果将图型左右交换一下（即用绿片在左，红片在右），则可以看到抛物线是右近左远，这时图型的频谱也是左右交换一下，即左边的高峰向低频移，右边的向高频移。而人的心理物理实验也说明了如果左边是低频条纹，右边是高频条纹，则融合起来的“栏杆”就是右边离人近左边离人远。因此说明了上述工作和人的心理物理实验吻合的情况；同时也说明了视差和空间频谱之间是具有一定的关系。

至于频谱上这些高峰的平移究竟说明了什么？我们知道物平面上图象的平移，其空间频谱上振幅频谱不变只是相位有变化，而物平面

上的相位变化表现在频谱平面上则是振幅频谱平移。现在一个三维物体在左右两眼的频谱高峰有位移，从定性上看来（因频谱细节不完全相同）可以说在物平面上的物象已包含有相位的信息了。而在本文所加工的左右两眼的物象只具有视差，因此是否可以这样认为：视差就反映了相位的因素了呢？这是一个很有意义的问题，是需要进一步加以研究的。

总之，从我们所得的结果中，我们认为：

1. 具有视差的立体图对的空间频谱高峰有相对的移动；
2. 移动的偏向是和三维物体离观察者的远

（下转第 41 页）

米远处听到。为了防止海狮在海里自由猎食,给海狮带上泡沫尼龙制成的口罩。另外,给海狮设计了一个特殊的取物钳,戴在海狮的嘴上。海狮带着取物钳潜入水下,找到火箭模型,并对准火箭的特定部位,把钳子夹在火箭上。然后,水面工作船利用与取物钳相连接的长绳把火箭拉出水面。

训练工作由浅水渐向深水发展。经过 15 个月的训练,海狮用取物钳取回物体的深度已达 152 米。

3. 虎鲸和领航鲸

训练虎鲸和领航鲸的方法和训练海狮相同,目的也是考察它们从海底取回物体的能力。为了防止它们逃跑,在它们身上装有无线电发报机,天线从背鳍前方伸出。如果它们逃跑,可以进行追踪。

从训练的结果来看,领航鲸比虎鲸更驯服一些。虎鲸利用取物钳取回物体的深度是 259 米,时间是 7 分 40 秒。而领航鲸携带取物钳已潜到 504 米深处,时间是 12 分 30 秒。如果不带取物钳,潜水深度可达 600 米。

另外,对领航鲸还作过一次新的实验。这次给领航鲸设计了一个新的取物钳,它带有一个由胼气发生器组成的自动起浮装置。只要领航鲸把取物钳与水下物体结合,就能触发胼气发生器。胼气发生器产生的大量胼气把气球吹胀,使水下物体浮上水面。虽然这种自动起浮

装置增加了取物钳的重量,但领航鲸还是接受了。领航鲸在一次实验中,利用它把一个沉落在 9 米深海底的鱼雷模型浮上水面。胼气发生器的声音和气球的膨胀等,都没有使领航鲸害怕。相反,它曾几次好奇地用嘴去碰碰气球。

从这些实验来看,海豚、海狮、虎鲸和领航鲸经过训练都能完成某些海上任务。将来,这些动物也许会成为海上“警犬”帮助人们工作。

六、鲸类与海洋调查

大多数鲸是回游的,有些海豚也回游。

过去曾用“标签放流法”来跟踪它们的回游路线,即用金属标签固定在它们身上,然后释放,待渔民再次捕到它们时,就可以知道它们某时到达某地。但是标签放流的回收率是相当低的。现在已开始用无线电发报机来跟踪鲸和海豚的回游路线。小型无线电发报机固定在鲸类的背部,天线从背鳍前方伸出。

海洋工作者经过多年的试验,现在已初步试验成功用鲸类来遥测海洋的物理参数,如海流、温度、盐度、含氧量、海水的声学性质等。测量方法是这样,把小型测量仪器和无线电发报机一起装在鲸类的身上。测量仪器测得物理参数由发报机发出。发报机发出的信号还可以由卫星转播。这样,把鲸类发报与卫星系统结合起来,为海洋调查提供了新方法。

(上接第 30 页)

Von Krebsen und Insekten. Leipzig-Wien: F. Deuticke, 1891.

[3] Burtt, E. T. et al: *Proc. R. Soc. B*, **157**, 53, 1962.

[4] Burtt, E. T. et al: *Symp. Soc. exp. Biol.* **16**, 72,

1962.

[5] Allen, J. L.: The optical functioning of the superposition eye of a nocturnal moth. Thesis, Massachusetts Institute of Technology, 1968.

(上接第 34 页)

近有关的。如果是一个单调变化的物体,则离人近的一侧频谱高峰向高频移,离人远的一侧频谱高峰向低频移;

3. 从定性的角度来看,这个平移现象是说明了两个物象已经包含了有相位信息。

参 考 资 料

[1] Julesz, B.: *Scientific Amer.*, **212** (2), 38, 1965.

[2] Barlow, H. B. et al: *J. Physiol. (Lond.)*, **193** (2), 327, 1967.

[3] Hubel, D. H. et al: *Nature (Lond.)*, **225**, 41, 1970.

[4] 光学信息处理简介,光学机械,1972 年 5 月第 39 页。

[5] Campbell, F. W. et al: *J. Physiol. (Lond.)*, **197**, 551, 1968.

[6] Maffei, L. et al: *Vision Res.*, **13**, 1255, 1973.

[7] Enroth-Cugell, C. et al: *J. Physiol.*, **187**, 517, 1966.

[8] Fiorentini, A. et al: *Vision Res.*, **11**, 1299, 1971.

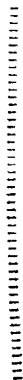


图 5a 有规则条纹的质地图

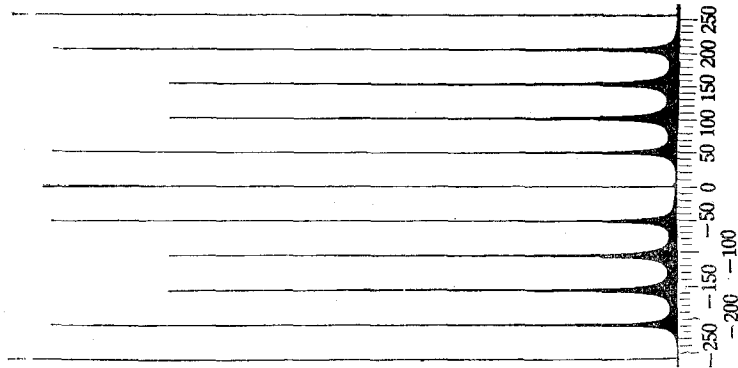


图 5b 上图的频谱

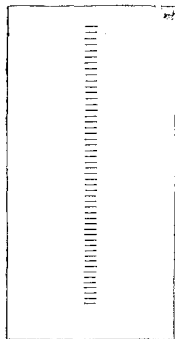


图 5a 左 左眼的投影



图 5b 左 上图的频谱

有规则条纹质地的抛物线

$$\begin{cases} x^2 + y^2 = 100^2(x + 10). \\ y = 0. \quad -500 \leq x < 0. \end{cases}$$

在左右两眼的投影及其频谱

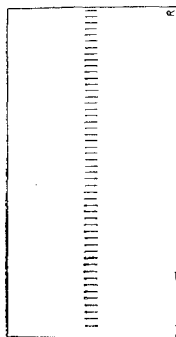


图 5a 右 右眼的投影

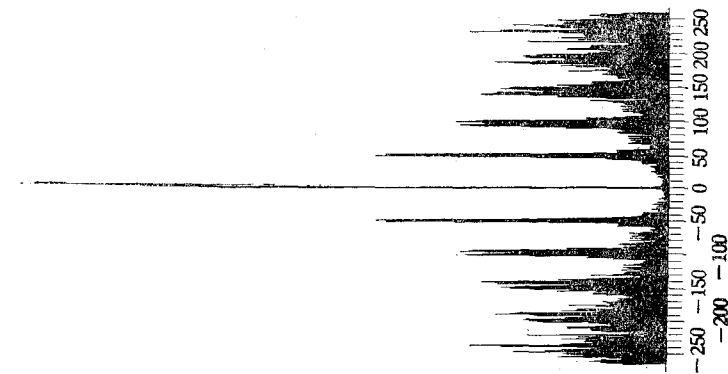


图 5b 右 上图的频谱

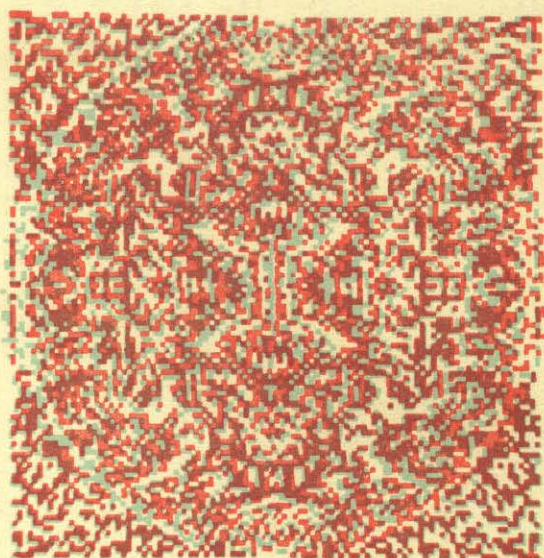


图3 椭圆抛物面 $x^2 + y^2 = 10^4(x + 10)$ 的立体图对

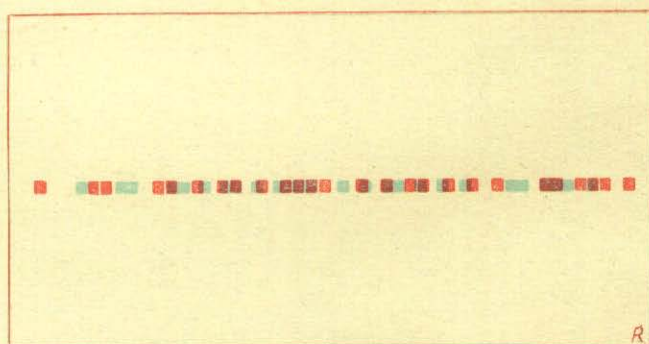


图6 套色的随机点质地的抛物线的立体图对 (即 4a 左, 4a 右)

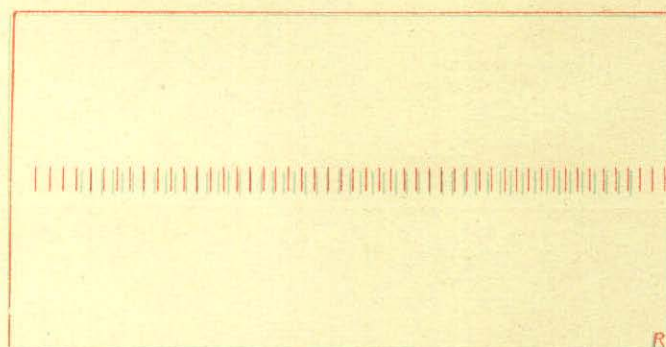


图7 套色的有规则条纹质地的抛物线的立体图对 (即 5a 左, 5a 右)