

维特根斯坦的“感知—语义连续论”和机器视觉研究

徐英瑾

摘要：“机器视觉”这个名目，可用于称谓各种以图像处理为基础的自动监视和机器人引导的技术。现在机器视觉已经发展为一个庞大的家族，其中每一个技术品种都为各自不同的技术目的而被开发。因此，还很难说有什么统一的科学基础将林林总总的机器视觉研究真正整合到了一起。然而，任何一个整合的机器视觉路径都必定会预设此种或彼种的哲学假设，而哲学家的任务就是去揭示之、检验之、批评之或者辩护之。在这里，被牵涉到的最核心的哲学争议乃是：视觉模块处理输入信息的组织原则，是否和中央推理系统的信息处理原则相互连续？抑或彼此分殊？以维特根斯坦的知觉理论为出发点，可以为“连续性论题”做出辩护，并由此为机器视觉的研究提供一条可能的思路。

关键词：机器视觉；感觉材料；视知觉；维特根斯坦；格式塔原则

中图分类号：B561.59 文献标志码：A 文章编号：2095-0047(2012)04-0133-15

一、导论

所谓“机器视觉”(machine vision)或“计算机视觉”(computer vision)，其实质就是用工程学方法对人工感觉系统所提取的视觉信息进行提取和读解，从中解析出关于所视对象的关键信息(如颜色及三维形状等)，以便人类用户或人工智能系统的中央语义处理器对其给予更为深入的处理。不难想见，在工业探伤、自动焊接、邮政自动化、导弹制导、图像自动解释、高水平的人机活动等应用领域，具有一定灵活性和可学习性的机器视觉系统所可能创造的社会效应乃是不可估量的。但就目前已经出现的机器视觉系统而言，它们在背景噪音分离、三维成像等诸方面，离人类视觉系统的水准还很远，因此依然有很大的改进空间。

作者简介：徐英瑾，复旦大学哲学学院副教授。

基金项目：本文受复旦大学“985工程”三期整体推进人文学科研究项目(项目编号：2011RWXKZD013)以及霍英东教育基金会第十二届高等院校青年教师基金基础性研究课题(项目编号：121095)的资助。

目前主流的机器视觉研究,基本遵循如下的模式。依据英国视觉专家大卫·马尔(David Marr)的“视觉阶段三分法”^①,视觉问题分为“初等层次视觉”、“中等层次视觉”和“高等层次视觉”三个子问题,并在这三个子问题下附设更多的子问题。每个子问题或子—子问题都具有相对的独立性,其解决都依赖于一些非常特殊而专门的程序设计。总而言之,我们既看不到这些专门的技巧设置之间的理论联系,也看不到所有的这些构建和智能体的中央语义处理系统之间的关联。对于此弊,人工视觉方面的国际权威戴维斯(E. R. Davies)教授曾给出过如下评论:

在本书中,我们已经讨论了如何处理图像以消除噪点,如何检测特征,如何通过对象的特征而为其定位,如何建立光照模式,如何为自动视觉监视设计硬件系统等等。对于本话题的研究已经延续了四十年来,而且一路以来我们确实已经经历很多。但这项研究依然是零敲碎打的,而不是富有系统的。通常的情形是,技术的发展由小团体中的科研人员的特殊兴趣所激发,而其成果也往往只是为特定问题量身定做的。与此情形相应的,则是如下事实:各种算法、程序、技巧都受限于不同研究者各自的创造性。^②

很显然,对于机器视觉问题的研究需要某种宏观的哲学视角,而工程学层面上的零敲碎打是于事无补的。本文对于视觉问题的讨论,就将以后期维特根斯坦的哲学为指导。众所周知,后期维氏哲学的核心命意,是对日常生活中的“语言游戏”规则予以澄清,并由此驱散千百年来形成的各种哲学迷雾。从表面上看来,这和我们在此所关心的视觉机制问题,似无明显关涉。但只要我们仔细阅读他在《哲学研究》第二部分、《关于心理学哲学的评论》以及《关于颜色的评论》中的相关议论,我们就不难发现,如何从哲学角度理解视知觉机制的工作方式,乃是他后期思想中的一个非常重要的话题,其中蕴藏着大量的思想宝藏,亟待学界予以系统地整理和阐发。

二、维特根斯坦关于视觉问题的核心论点:“感知—语义连续论”

在笔者看来,维特根斯坦关于视觉问题的核心论点,可以被概括为“感知—语义连续论”。“感知—语义连续论”的具体内容为,我们和真实世界的信息交换固然需要来自于世界的物理刺激来催生初级的感觉材料。但是,这些感觉材料的构成形式,并不与中央语义处理系统的语义构成规则是彼此异质的。换言之,在他看来,高层的语

^① David Marr, *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*, New York: W. H. Freeman and Company, 1982.

^② E. R. Davies, *Machine Vision: Theory, Algorithms, Practicalities* (Third Edition), Singapore: Elsevier, 2005, p. 833.

义规则能够随时渗入这些感觉材料(尽管实际上未必时时渗入),感觉和语义推理之间的关系乃是连续的,而非断裂的。

现在我们就结合维特根斯坦本人的哲学文本,对这一核心论点加以详细说明。

从维氏自身的思想嬗变角度来考察,主张“语义认知—感觉连续论”,乃是其知觉理论的一贯特色,此特色甚至在其《逻辑哲学论》阶段就已初露端倪。在该书的论题 2.0131 中他曾提到,视野中的一个斑点虽未必一定是红色,却肯定具有某种颜色,“可以说它是被一个颜色空间所包围着的”^①。这里所说的“颜色空间”,显然是指某种语义学空间,而对此段评论的蕴意也可以做如下解释。认知系统对一个被输入的感觉刺激的把握,必然会伴随着一个高层的、关于颜色概念的语义网络的介入;认知系统可能会在判断事物的颜色方面出错,但在判断过程中,该颜色语义网络的介入却始终是一个必然条件(没有这个条件,甚至谈不上颜色判断的发生)。

此观点在维氏后期哲学中得到了更为系统的发挥。在其遗稿《关于颜色的评论》开首,他就提到了两类语言游戏,一类是呈报“某物的颜色比另一物更暗些或更明些”,另一类则是呈报“两个色调之间的明暗度之间的关系”。维氏很清楚地看到,呈报色调关系的语言游戏在实质上和前者非常不同。后者牵涉到某种非时间的“内在关系”,而前者则牵涉到了某种具有时间性的“外在关系”。^②再套用今日认知心理学的术语来说,所谓“非时间的内在关系”,指的就是(关于颜色的)语义网内部各概念节点之间的拓扑学关系——这种关系具有一定的稳定性(stability)或鲁棒性(robustness),并在一定意义上和系统实时获取的感觉材料相脱节。而所谓的“带有时间性的外在关系”,则是指原始的感觉刺激和相关颜色概念节点之间的刺激—反应关系,这种关系的建立具有明显的实时性,其特征恰恰取决于怎样的感觉材料在当下环境中被系统所获取。由此不难想见,既然维氏将语义节点之间的关系看成是内在而稳定的,并将感觉和语义节点之间的关系看成是外在而可变的,那么他也就应当能够同情那种“自上而下”(或至少是“上下兼修”)的知觉重构方案,即仅仅把感觉材料看成是激活内部语义网络相关节点的刺激物,而并不是将其视为知觉活动的本质性因素。

而由此带来的理论后果,显然就是对符号语义层面上的心理认知活动之独立性的强调,或换言之,在维氏看来,我们对于关于颜色的语义网的心理学知识,根本就不需要被还原为我们关于视觉神经系统的生理学以及物理学知识。维氏在《哲学评论》§218 中写道:“我们需要的是一种心理学的甚或说是现象学的颜色理论……(在该理论中)不出现假设的物体,如波长,细胞等等。”^③从这个角度看,虽然维特根斯坦

① Ludwig Wittgenstein, *Tractatus Logico-Philosophicus*, trans. by D. F. Pears & B. F. McGuinness, London: Routledge & Kegan Paul, 1961, p. 6.

② 参见 Ludwig Wittgenstein, *Remarks on Colour*, trans. by Linda L. McAlister & Margarete Schättle, Malden: Blackwell, 1977, p. 2.

③ Ludwig Wittgenstein, *Philosophische Bemerkungen*, Frankfurt am Main: Suhrkamp Verlag, 1984, S. 273.

没有机会读到日后马尔的视知觉加工理论,以及邱琪兰德(Paul Churchland)关于“神经语义学”的构建工作^①,他应当是不会赞成那种针对视知觉过程的“自下而上”的重构方案的。

然而,为何维氏主张的这种“连续论”就一定就是对的呢?我们是否能够从维氏文稿中找到相关论证,以支持这种“连续论”呢?笔者从《哲学研究》对于“两可图”的大段讨论中^②重构出了如下归谬论证:

1. 如果视觉信息的加工过程所遵循的原则迥异于语言规则的话,那么我们就一定能够找到这样一些心理学对象,它们的出现能够被认知主体意识到,但是认知主体的更为高层面的语义知识却无法渗入它们(这是归谬论证的前提)。

2. 现在我们把上述一般前提予以推广。如果“感知—语义断裂论”是对的,那么,若将其套用到“两可图”的例子上来,我们就必须承认,在认知主体审视“兔鸭图”的时候,无论其看出这是一只鸭,或只是一只兔,抑或看出它既可被视为鸭,亦可被视为兔,主体所获得的最底层的的感觉材料乃是一样的。他对图片的不同的语义解读,不会影响其所获取的视觉印象的最初组织形式。

3. 但这一不变的最初组织形式存在的证据又是什么呢?我们能够发现的,恰恰是主体在审视这张图片时视觉组织形式的变化,比如心理学中的“眼动”测量实验会告诉我们,被试者的视线的扫视次序,会随着他自己对图片的语义解读方式的改变而改变。当他呈报自己读出鸭子的时候,眼动检测设备会发现他的视线是从兔—鸭头的鸭嘴部分开始向下扫描的。当他呈报自己读出兔子的时候,眼动检测设备会发现他的视线是从兔头部分开始从右往左扫描的。当他呈报自己同时读出鸭子和兔子的时候,眼动检测设备会发现他的前述两种视线扫描路径现在已经混作一片,交相竞争。总之,没有任何一种扫描形式可以不受制于语义解读方式的变更而保持自身的不变。^③

^① Paul Churchland, “Neurosemantics: On the Mapping of Minds and the Portrayal of Worlds”, in *The Emergence of the Mind: Proceedings of the International Symposium*, edited by K.E. White, Milan: Montedison the Fondazione Carlo Erba, pp. 117–147.

^② 参见同上书,第202页以下对于“看见”的哲学语法的讨论。

^③ 更技术性地来说,眼动(eye movements)由以下四个类型构成:视扫(saccades)、平滑追踪(smooth pursuit)、平稳定位(stabilized fixation)以及双眼聚散(vergence)。眼动对于人类获取信息来说,具有以下两个重要意义:第一,人类的精确视觉的获取区域——视网膜中央窝(fovea)——其实是很小的,为了获得高像素的图像,眼球就必须不断移动,使得对象的各个部分都能够被中央窝所聚焦;第二,人眼的成像速度其实不快,锥状细胞得花费20毫秒才能够完成对光亮刺激的一个充分生化反应。为了在这段时间内保持中央窝和对象之间位置的相对稳定(以防止图像糊掉),眼动机制就必须保证眼球能够“盯住”相关对象。但这种“盯住”本身是动态的,因为在此过程中,可能人的身体和物体都在做一种绝对运动。对于眼动机制的深入科学探索,请参见 Micheal F. Land, Benjamin W. Talter, *Looking and Acting: Vision and Eye Movements in Natural Behaviour*, New York: Oxford University Press, 2009.

维特根斯坦本人没有明确地提到眼动实验,但在《心理学哲学评论》§23–24中,他间接提到了这个话题,请参见 Ludwig Wittgenstein, *Bemerkungen über die Philosophie der Psychologie & Letzte Schriften über die Philosophie der Psychologie*, Frankfurt am Main: Suhrkamp Verlag, 1984, S. 23。在那里,他提到了人们(下转第137页)

4. “断裂论”者或许会反驳说,这一不变的最初组织形式实际上是如此之基本和自我封闭,以至于其可以逃过任何外部物理设备的检测——包括眼动检测设备的法眼。但问题是,既然这个层面的存在难以找到物理证据,那么相信其不存在和存在就不会带来任何理论后效上的分别。换言之,设定其存在不会给我们对于视知觉过程的解释带来任何理论便利(奥卡姆剃刀原则)。

5. “断裂论”者或许还会反驳说,这一不变的最初组织形式实际上只能够通过内省方式来发现,而内省活动本身则在原则上不会留下任何物理证据。但问题是,在维氏看来,即使我们退一步接受了内省原则,我们也无法确定这个稳定的感官层(或所谓的“内部图像”)的存在。这是因为,“‘内部图像’这个讲法本身就是误导人的,因为它把一个‘外部图像’当做了自己的模板”^①。换言之,“纯粹感觉”只是我们在受到语言误导以后臆想出来的一个和“外部图像”对应的镜像而已,其存在并无直接的“现象”证据。由此看来,当我“发现”兔—鸭图中的两个不同面相指向的是同一个对象的时候,这同一个对象也可以被理解为引起感觉的物理对象,而不是被物理对象所引起的感觉。

6. 所以,对那个不可被语义渗透的感觉层面的设定,就无异于“盒子里的甲壳虫”,完全可以在理论构建中被“约减”^②。

7. 所以,“断裂论”是错的。

8. 所以,对于视知觉进程的重构方案,更应当采纳“连续论”的进路,即承认高层语义对于低层感觉具有强大的穿透力。

这是笔者从维氏的零碎哲学评论中所能够重构出来的最有力的论证。此外,根据笔者的理解,维氏关于视知觉机制的“连续论”还附带了三个重要的延伸性论点。下文将对其一予以展示。

三、延伸性论点(甲):“低级知觉—高级知觉”之分, 并非“无语义—有语义”之分

“感知—语义断裂论”的支持者,或许会以所谓“低级知觉”和“高级知觉”之间的

(上接第 136 页)看待同一个三角形的不同方式(在一个方式中将这个角视为顶角,在另一个方式中将另一个角视为顶角)。他也提到了,在不同的扫视三角形的方式中,主体的视线扫描路线是彼此分殊的。同时他还强调,这种视线扫描方式的形成并不是主体刻意解释的产物,因此这只能是“看”的结果而不是“思”的结果(这里需要注意的是,他说的“看”并不是一种纯粹的感官性的看,本身还是带有最起码的语义内容的)。从他的这段评论来看,维氏应当是赞成这样一个观点的:在反思性的意识活动还没有被充分带入自觉意识状态的知觉状态中,某种微妙的身体活动(如眼动)其实已经被预先激活。因此,前反思的知觉活动之间的类型差异,就会通过伴随性的身体活动之间的差异体现出来。

^① Ludwig Wittgenstein, *Philosophical Investigations*, p. 206.

^② “甲壳虫”之喻采自《哲学研究》§293, 参见 Ludwig Wittgenstein, *Philosophical Investigations*, p. 106.

差别来为自己的观点辩护。换言之,在他们看来,语义可以渗入“高级知觉”,却无法进入“低级知觉”。维氏则明显不同意这个观点。他承认,在我们的不同种类的知觉活动之间,固然存在着复杂程度方面的明显区分,但这个区分却并非“无语义介入的信息加工过程”和“有语义介入的信息加工过程”之间的区分。一个很简单的例子就能够说明这种分别。假设现在有一只兔子出现在张三、李四和王五跟前,而这三人则有三种回应方式:

(1) 张三惊喊道:“兔子啊!”

(2)(听到张三惊呼的)李四则呈报说:“我也看到了兔子。”

(3)(只来得及看到兔子尾巴的)王五则呈报说:“我将其视为兔子。”

关于(1)和(2)之间的区别,维氏写道:“呈报和惊呼都是对知觉和视觉体验的表达。不过,惊呼是一种和报告意思不同的表达,它从我们身上挣脱出来——它与体验的关系,类似于喊叫和疼痛的关系。”^①用认知心理学的术语转译来说,在“惊呼”的过程中,感觉体验和语义网中特定的语义节点(如“兔子”)建立了相当明确的联系,这就使得系统能够以较低的资源消耗成本,快速地对相关体验做出语义明确的语义描述(不过,这种过于简单的语义介入方式也造成了一种假相,即“惊呼”是一种“挣脱”于身体的直接体验,而与语义认知无涉)。与之作对比,在(2)所描述的“呈报”过程中,语义介入的模式却在相当程度上被复杂化了,或用维氏自己的话说:“如果你产生了由惊呼表达的视觉体验,那你就在思考你之所见”。很显然,一个认知系统对其所见的“思考”,恰恰意味着更大范围内的语义网节点被牵涉到了对于当下感觉材料的评估过程之中——换言之,知觉活动中原先“潜伏”的语义认知因素,现在被全面凸显了出来。

而一旦这种凸显程度跨过一定的阈值,那么(2)所描述的那种认知模式就会被转变为(3)。有意思的是,维氏似乎很在意(2)和(3)之间的差异,并时时强调之。他把二者之间的区分称为“看”(sehen)和“看作”(sehen...als)之间的区分,并认为使用这两个词(组)的语言游戏彼此非常不同。他写道:

如果我说“我把它看作……”,这对我来说,就像我在看到刀叉时“我现在将其视为刀叉”一样没有意义。这样的表达不会被人理解——就像“它现在在我看来是一把叉子”或“它也可以是一把叉子”不会被人理解一样。^②

维氏的言下之意是,以“看作”为关键词的那类语言游戏,其启动门槛要高于以“看”为关键词的那类语言游戏。具体而言,除非确有其他的可能性,使得某人会将眼

^① Ludwig Wittgenstein, *Philosophical Investigations*, p. 207.

^② *Ibid.*, p. 205.

前的某物不看作这而看作那,否则他就将无法合法地使用“看作”这个动词词组,而只能用“看”这个动词。若换用我们今天的话语体系,这也就是说,在一个认知系统将物理世界中的物理对象甲(如兔—鸭图)“看作”语义对象乙(如“兔”)的时候,那么,除非发生下述情况——至少有一个别的语义对象丙(如“鸭”)能够像乙一样,恰如其分地完成对于甲的语义解读,且乙和丙彼此不同义——否则,这个动作就仅仅是“看”而非“看作”。

现在我们可以就讨论得再深一点。从信息加工的角度来看,当系统将甲“看作”乙或丙的时候,其中央信息处理中枢实质上正在进行一系列异常复杂的信息处理活动。比如,或许有同样权重的不同证据体在支持“这是鸭子”和“这是兔子”这两个彼此冲突的判断,而系统正在努力寻找新的证据体来将这两个判断中的一个予以保留,一个予以删除(否则系统就难以维护知识体起码的自治性)。如果因为证据来源不足而使得系统无法对这两个判断作出取舍的话(这是经常会发生的情形),那么系统或许就会因此产生一些负面的情绪反应(如困惑、焦躁、忧虑等等)。很显然,所有这些复杂的信息加工活动,将使得更大范围的语义网节点被激活,或更多的系统资源被消耗,甚或在某些极端情况下,导致系统着手修正原有语义网的拓扑学结构。一言以蔽之,如果说“看”这个动作主要牵涉到的是在局域语义网指导下的简单语义分配过程的话,那么“看作”这一动作则无疑向我们暗示了,在上述语义指派进程产生不融贯的知识体之时,智能系统为消除这种不融贯性,必须做出某种更大的努力(至于这种努力是否能有所成果,暂且不论)。也正是基于此,在维氏看来,“看作”实际上并非“知觉”的一部分,而是“思考”的一种形式。^①

但正如前面所反复提及的那样,维氏心目中“看”和“看作”的区分,依然只是两种语义介入模式之间的区分,其关系类似于“自动导航状态”和“主动导航状态”之间的区分,却绝不是“无导航状态”和“有导航状态”之间的区分。因此,他对于该区分的阐述,非但没有违背,甚至还在一个方面丰富了其知觉理论的核心命意,即语义认知的高层引导机制对于感觉印象的构成而言,始终具有关键意义,而无论这种引导机制本身是否被更为高层的心智活动所反思到。

四、延伸性论点(乙):感觉材料组织方式之“完型”特征, 乃是语义推理之灵活性的体现






熟悉心理学史的读者可能会指出,前面所反复提及的高层语义认知对于低层感

^① Ludwig Wittgenstein, *Philosophical Investigations*, p. 207.

觉的渗透性,绝非维氏之独家论点。格式塔心理学(或“完型心理学”)也认为,视觉意象的构成需要一些“自上而下”的组织原则予以牵导,因此,在他们看来,对于无组织形式的个别感觉原子的逐个认知其实是不可能的(换言之,不存在着高层组织形式无法渗入的感觉原子)。在这个问题上,维氏知觉理论和格式塔心理学之间又有何实质差异呢?

一种最粗略的比较或许会告诉我们,格式塔心理学家总是乐于概括出一些人类知觉系统借以进行视觉印象组织的一般性原则,并使得其论述具有比较强的系统性(见下表所示)。至于维氏的讨论,虽在一定程度上引用了格式塔心理学的心理学实验案例,但在论述形式上却显得相对散漫。换言之,他似乎并不屑于将视觉组织的基本原则加以罗列和归纳,而仅仅满足于向读者展现和视知觉相关的种种语言游戏的丰富性。

表 1 诸格式塔原则简表^①

诸格式塔原则的名称	原则的内容	例证以及相关说明
图像—背景原则	在知觉主体的视野中,若一个物体(图像)看起来是具有凸显性的,那么在视野中的其它东西就会消退到背景中。	 此图即可被视为两黑色的人类侧影之相向而对,亦可被视为一白瓶在黑色背景中凸显自身。如何知觉,取决于何物被凸显:侧影还是杯子?
邻近性原则	在对被知觉到的东西进行归类时,易于把相近的东西视为一组。	 在上图中,我们倾向于将中间四点看成两个组。
相似性原则	人们倾向于把彼此相似的事物归为一类。	 在左图中,我们更倾向于将其视为四列空心圆和实心圆,而非四行空心圆和实心圆。
连续性原则	人们倾向于知觉到连贯的或者连续流动的东西,而非断裂的或不连续的形式。	 此图很容易被视作一条连续曲线和连贯直线的相交图。若将其视为一个“V”字和一个倒“V”字的上下叠加,则显得不太自然。
闭合性原则	人们倾向于把并非完整或闭合的图形看成是完整的或者闭合的。	 此图很容易被视作三角形,尽管实际上这“三角形”的每条边都是断裂的。
对称性原则	人们倾向于把事物都知觉为沿中心左右对称的样子。	{[]}<0> 在此图中,我们明显知觉到了四组括号,而不是八个彼此无关的单括号。

^① 改写自斯滕伯格:《认知心理学》,杨炳钧、陈燕、邹枝玲译,北京:中国轻工业出版社 2006 年版,第 93 页表 4.3。

但这种论述风格上的区分毕竟还是浮面上的。我们不妨把问题想得再深一点。前文已反复提及,维氏心目中的视觉印象组织规则,在根底上实为关于视知觉的语言游戏规则,因此,他很可能将视觉中枢的信息加工规则看成是语言中枢的信息加工规则的某种变异,或将其视为后者的外围衍生物,或干脆认为这两者本来就分享着完全相同的信息加工原则。但我们也都知道,语言中枢的信息加工规则本身就具有很大的伸缩性和弹性,因为不同的经验主体由于各自不同的学习经历,很可能会拥有彼此不同的信息加工规则(如色盲患者对于颜色刺激的信息加工规则,就肯定不同于正常人)。这也就是说,我们至多只能对这些规则进行一种“元叙述”(比如说“这些规则具有很大的伸缩性”本身,就是一种元叙述),而无法给出一种具有足够普遍性的“对象性叙述”。而在维氏的知觉理论框架中,由于语言游戏规则和视觉组织原则之间的同质性,对前者的“元叙述”所具有的这种伸缩性,自然很快就会传递给对后者的“元叙述”。或用他自己的话来说:“‘对被看见者的表征’这个概念,就像‘复制品’这个概念一样,具有很大的伸缩性,‘被看见者’这个概念也是如此。这两个概念密切相关(但这并不是说它们很相似)。”^①

在今日的心理学界,也有人采纳和维氏相近的论点,即认为高级语义认知(“想”)的介入,对于知觉把握(“看”)的形成具有巨大的“塑型”作用。美国心理学家罗克(Irvin Rock),就是此种见解的主要倡导者。对于这种高级语义介入能力的强调,还构成了一派独特的知觉理论,叫“建构知觉”或“智能知觉”,其核心论点是:

在建构知觉中,知觉者以感觉信息为基础,并利用其他信息源,对一个刺激建立或建构一个认知理解(知觉)。这种知觉又叫智能知觉,因为该理论指出,高级思维在知觉中起到了重要作用。^②

或用罗克本人的话来说:

尽管和更高阶层的心智能力——如有意识的思维,以及对被意识到的知识的运用——相比,知觉似乎是“匿名”的,但我还是想说,知觉在实质上乃是具有智能的。之所以说它是“智能”的,是因为知觉乃是奠基在和思维类似的种种心灵进程之上的,这些进程包括描述、推理和问题求解等——尽管它们乃是稍纵即逝的、无意识的、不诉诸于语词的。^③

① Ludwig Wittgenstein, *Philosophical Investigations*, p. 208.

② 斯滕伯格:《认知心理学》,第103页。

③ Irvin Rock, *Perception*, New York: W. H. Freeman and Company, 1984, p. 234.

罗克的上述评论或许会让维特根斯坦主义者感到“心有戚戚焉”。从某种意义上说,维氏的视知觉理论就可被视为罗克“智能知觉”论的先声。

五、延伸性论点(丙):三维视知觉对象的建立有赖于 身体运动模式的参与

维氏视知觉理论所带来的理论效应,不仅体现于对思、感壁垒之消弭,甚至还体现于对视觉和身体自我定位机制之间内在关联的揭示。换言之,在维氏看来,对于三维空间对象的大小远近的把握,并非视感觉模块一家之功,还需要身体运动模式的全面参与。

他的这个观点是分以下两个环节来阐述的。

第一,我们关于三维空间对象的大小远近的知识,不是在视感觉模块所提供的原始信息中被给予的,而是来自于后天的经验。

关于视感觉模块所提供的原始信息,维氏曾在20世纪30年代写就的《大打字稿》中以专门名目称之为,即“视觉空间”。而那种整合了三维空间对象的视觉信息,则相应地被他称为“物理空间”。有趣的是,这一区分在《哲学研究》以及以后的维氏著述中不再反复出现,这很可能是因为20世纪40年代以后的维氏,已不再相信存在着高阶语义知识不可渗入的纯二维感官信息。但他在30年代所做出的关于“视觉空间”的一个否定性判断——“若有视觉空间,则其无法提供关于三维空间对象的大小远近的知识”——却依然在其后期哲学中生效,因为这个判断一方面并未断言视觉空间本身定然存在,另一方面也可以间接支援后期维氏所秉承的“思—感连续论”(因为上述判断实际上就已经将关于远近大小的知识获取责任归诸视感觉模块以外的智能模块)。

在《大打字稿》中,维氏是通过这样一个例子来阐明上述论点的:

到底有什么因素去驱使我去解释:为什么我通过我的窗户所看见的这棵树要比窗户本身来得大一点呢?这得取决于我如何使用语词“较大的”与“较小的”——让我们想想在日常的/正常的/生活中,如下视觉经验究竟会向我们呈现出什么样子吧:我们见到了一些长度不等的棍子,但其中的每根棍子都被刻度线划分为(可见的)彼此等长的部分。在这里,我们可不能引入关于语词“更长”或“更短”的双重用法。这是因为,在某些情况下,我们会将那些被分成更多部分的棍子说成是较长的棍子。^①

^① Ludwig Wittgenstein, *The Big Typescript* (TS213, German-English Scholar's Edition), edited and translated by C. Grant Luckhardt and Maximilian A. E. Aue, U.K.: The Blackwell Publishing, 2005, p. 206. 另外请对照阅读 Ludwig Wittgenstein, *Philosophische Bemerkungen*, Frankfurt am Main: Suhrkamp Verlag, 1984, S. 101(§72)。

按照这段话的意思,如果缺乏对于眼与物之间的间距的经验知识(以及对于物体自身的大小的经验知识),那么,那种与物理意义上的“远/近”概念相互联系的“大/小”概念就不会自动呈现给智能体。比如,我们之所以知道一根棍子的实际长度并不是它直接呈现在视野中的长度,就是因为我们已经在事先掌握了通过计量棍子的等分刻度来估算其长度的物理方法,并且知道这种刻度等分系统本身是不会因为物体与观察者的眼睛之间的间距的变化而失效的。反过来看,在纯粹被给予的“视觉空间”中(如果它的确存在的话),什么东西被呈现为“大”的,它就是“大”的,我们并不去追问它实际上到底大不大。

维特根斯坦的这层意思,还可以通过一种带有科学哲学色彩的话语方式得到转述。如果在视网膜所获得的视觉信息是一种“被解释项”的话,那么就会存在着不同的“解释项”或“假设”以便说明其语义内涵,而且这些说明关系中的每一种都可以孤立地成立(比如,孤立地看,我们既可以将视觉空间中的短棍解读为长棍,也可以解读为短棍)。我们的知觉系统要在如此多的彼此竞争的假设性解释中予以甄别和遴选,就需要一个更为广泛的背景知识作为参考,而这个背景知识只可能来自于感觉模块之外。感觉模块本身则是无力完成对于三维空间的语义构建的。

第二,在三维对象的空间信息的获取方面,身体的运动模式所扮演的角色举足轻重。

现在又冒出了一个新问题:上面所说的“背景知识”又如何形成呢?在维氏看来,我们关于空间的背景性语义知识并不仅仅体现为一系列静态命题所构成的僵死知识体。这些知识可以随时更新、删改或扩充,而做出这些变更的关键环节便是身体。身体在连接感官和高层语义知识之间的信息渠道中扮演了一个非常微妙的角色。一方面,它为感官的物理存在提供了真实的空间基础。另一方面,它又是语义层所发出的语义命令(如“左转看墙壁”,甚至细微的眼动轨迹改变)的直接执行者。在上面举的看棍子长短的例子中,身体的运动模式所扮演的功能角色便可以大略地分析为如下过程:

1. 接受到视网膜上的信息(关于两根棍子的视觉形象)后,中央语义系统或许会产生一系列彼此差异的三维语义解释。

2. 中央语义系统启动假设验证机制以筛选这些解释。这会进一步激发身体改变自身的物理位置,并由此改变感官的观察位置,以期获得新的感觉材料。

3. 系统转而计算新的材料和每个假设的匹配度,匹配度最高的假设胜出。

4. 胜出假设进入背景知识,在以后的模式识别任务中不再需要检验,而可以被直接调用,除非遭遇到极强的反例。

从上述流程来看,如果我们把中央语义系统比作一位亚里士多德式的思辨家的话,那么身体就是其实验室——它使得亚里士多德式的自然哲学可以有机会进化为

伽利略式的实验科学。这又是因为,唯有身体才能够扮演“假设模拟”和“证据获取”这双重角色(见上述步骤第4步),没有其参与,中央语义系统产生的新假设就根本无从得到验证。有意思的是,关于这种身体无法执行中央语义系统的命令的极端情况,维氏还曾以以下一种新奇的方式予以了设想。

假设:我身体的所有部分都被分离了,直到仅仅只剩下一只眼珠。这眼珠被固定在某处不能动,且保持了能够看东西的能力。这样的话,世界又将向我显现为什么样子呢?我不能知觉到我自己的身体的任何一部分,并且我们假设:我的眼珠对于我而言是透明的,甚至在镜子中我也看不到自己的眼珠。现在就有了这样一个问题:通过我的视觉图像,我能进行自我定位吗?自然,“自我定位”(Mich lokalisieren)在此仅仅意味着去建立视觉空间的一种特定的结构。^①

这个“疯狂”的思想实验显然已经取消了我们通常意义上的身体所具有的两个特征。一是通过对于眼珠的固定安置剥夺了其可动性,二是通过对于眼珠的绝对透明性的假设来否定了肉体自身在视野中被呈现的可能。很显然,既然对于肉体本身的运作状态的任何观察都是不可能的,那么我们自然也就无法将肉身的运作把握为视觉现象变化的原因了,而恰恰就是这种把握,对于如何在客观世界中对我进行定位乃是必不可少的。在引文的最后一句话里,维氏很明确地把“自我定位”视为三维空间构造中的一个关键环节,因此,依据上述推理,身体的缺失也自然会导致智能体对三维空间的语义构建的失败。这个思想实验显然从反面说明了身体的运作模式在三维知觉构成中所起的重大作用。

那么,我们又当如何从哲学角度评价维氏关于身体运作和视知觉之间关系的这些阐述呢?有的读者或许会说,从哲学史既有的对于视知觉机制的讨论来看,强调高层语义推理机制对三维知觉对象构成所起的作用,并不算啥新鲜论点。英国哲学家贝克莱(George Berkeley)在1709年发表的《视觉新论》中早就指出,三维空间中体现出的深度(即知觉者和被知觉对象之间的距离)并不是感官直接给予的,而需要经由经验的后天联想作用,才得以在智能体的符号世界中渐渐成形。但在笔者看来,贝克莱所没有清楚明言的,恰恰是身体运动模式在后天的经验背景知识的形成过程中所扮演的角色——对此角色的必要性,他甚至间接地予以否认。在该书§49中,他刻意强调触觉对象和视觉对象并非彼此同一,二者之相连仅仅是联想约定使然。^②很显然,他的这种主观观念论色彩浓郁的主张,在客观上实际上已经否定了各个感官道之间信息的可交换性,而这又反过来使得中央语义系统所产生的任何一个关于视觉纵深的语义假设,都无法获得来自于触觉信息的确凿验证。一个按照如此知觉原则

^① Ludwig Wittgenstein, *Philosophische Bemerkungen*, S. 100–101 (§72).

^② George Berkeley, *Philosophical Writings*, edited by Desmond M. Clarke, Cambridge: Cambridge University Press, 2009, pp. 23–24.

进行空间构建的智能体,恐怕也不会有机会得到一个统一的、稳定的三维空间表征方式,并在此基础上和环境产生足够有效的交流。与之相比较,维氏视知觉理论恰恰强调了在构建三维语义空间的过程中身体其他官能和视觉官能的彼此印证关系,这就为一个统一的三维空间表征方式的形成扫清了理论障碍。

有的读者或许还会说,在当代西方哲学的范围内,就身体运作和视知觉之间的内在关联这个话题而言,法国哲学家梅洛—庞蒂的讨论(Merleau-Ponty)比维氏的相关阐述内容更为丰富和系统。对于这一点,笔者并不否认,但我还是希望提醒大家注意维氏“身体”观和梅氏身体观之间的本质不同。在梅氏的哲学框架中,他所说的“身体”乃是可以现象学意义上呈现出来的肉身,而不是生理学所研究的客观身体。而在我们前面所看到的维氏的那个关于“透明眼球”的思想实验中,他所说的身体却恰恰就是客观意义上的身体(而非现象学意义上的身体)。从这个角度看,两个人对“身体”的讨论实际上并无太大的交集。看得更深一点,从哲学论证的角度看,梅氏从现象学身体出发来讨论知觉的路数,颇有“乞题”(begging the question)之嫌疑。这是因为,如果现象学身体已经被定义为“在知觉活动中呈现出来的身体”的话,那么从这样的“身体”出发来建构一种知觉理论,就等于在前提中预设了自己需要论证的结论。很显然,维氏就身体和知觉问题之间关系的阐述,并没有预设身体本身具有现象学属性,因此完全就可以洗脱这一嫌疑。

六、维特根斯坦的视觉理论和关于主动视觉的工程学研究(代总结)

那么,维氏视知觉理论(尤其是其关于视觉—身体关系)的讨论,对于机器视觉的研究,又有何启发意义呢?

在本文开首处笔者已经说过,目前机器视觉研究的主流方式,就是把视觉进程切割为一个个小问题加以专门化的处理,而不问视觉机制和语义推理系统之间的宏观联系。这种做法在精神上显然是和维氏所主张的“感知—语义连续论”背道而驰的,却在某种意义上和他所反对的“感知—语义断裂论”暗通款曲。

一些比较富有哲学洞见的机器视觉专家,则多少意识到了这种“零敲碎打的”传统工作方式的缺憾,尽管这种认识也未必是得缘于他们对于维特根斯坦的阅读。在某种模糊的哲学直觉的引导下,他们开始追求某种意义上的技术路径整合,以求构造出更为有效的机器视觉系统。于是一个新的机器视觉流派在近年开始走红,这就是所谓的“主动视觉”(active vision)流派。关于“主动视觉”的内涵,戴维斯是如此解说的:

……主动视觉所意谓的,乃是如下两个技术路径之一:(1)对摄像设备加

以移动,以求聚焦于或施加注意力于当下我们所感兴趣的环境中的某事项;或(2)施加注意力于图像中的某事项或某区域,与之发生动态信息交流,直至获取一种可被接受之解释为止。无论在以上任何一种情况中,都存在着一个注意力的焦点,而且视觉系统总是以一种主动的方式花费其时间预算,追踪并分析在场景中所出现的东西。它既不泛泛地将场景解释为一个整体,也不试图记住所有的东西。另外,在主动视觉中,下述任务是交付给一个更高层的处理器来完成的:对每项兴趣本身进行识别,并提出如何满足它们的攻略要求,并最后回应这些要求。^①

从上述定义来看,主动视觉流派背后的基本想法似乎和维特根斯坦很相合。二者都赞同,对于三维空间的语义构建需要一个动态的假设验证过程,而且在这一过程中,我们都需要设法让智能体的感官(无论是肉眼还是摄像机)的空间位置发生自主性位移。二者似乎也都赞同,视觉的运作机制本身预设了一个更高阶层的中央语义处理系统的运作,这个系统本身将向视觉机制负责提供最初的语义假设,以及运作资源的分布方案,还有对假设的最终验证评估结果。

但更为深入的比较将告诉我们,主动视觉流派和维氏之间的这些相合仅仅是浮面上的。二者之间的实质性差异体现在:

第一,在维特根斯坦的知觉理论背后,有着一整套的关于智能体的知觉活动之实质的系统化看法,或说得更彻底一点,他的知觉理论,其实只是他关于人类语言活动的一般看法的理论衍生物。而主动视觉的哲学动机则相对单纯得多,他们主要关心的是,如何通过摄像机的自主位移来减少系统所要处理的原始信息的数量,以此提高系统的运作效率。具体而言,他们虽然也认识到主动视觉的工程学构造无法回避高级语义知识的渗入,并承认他们必须设定一个能够自主运行的身体,但是为了工程学上面的便利,他们其实还是通过别的办法回避了这两个最关键的理论节点(比如,一些研究者就通过人类用户语义知识的干预来回避第一个理论节点;另一些则通过把主动视觉设施设计为“人体可穿戴”的“外置机器人”来回避第二个理论节点^②)。但这种回避显然把关于视觉判别的最关键的环节交付给了人类用户,而不是人工系统。从这个角度来看,激发主动视觉派之工作的基本动机,与其说是理论性质的,还不如说是工程学性质的。

第二,主动视觉流派的工作和传统的人工视觉研究一样,一般都落实为零碎的子课题,缺乏宏观层面上的整合性。进而言之,这些子课题各自具有彼此不同的解决目标、理论预设以及技术实现手段,彼此之间的联系若有若无。与之相比较,维氏对于

^① E. R. Davies, *Machine Vision: Theory, Algorithms, Practicalities* (Third Edition), pp. 573-574.

^② 参见牛津大学工程系“主动视觉研究小组”官网, <http://www.robots.ox.ac.uk/~lav/>。

人类视知觉问题的不同评论之间却有着明显的内在联系,并都服从于统一的哲学原则。从这个角度看,一种更忠实于维氏思想的人工视觉的工程学研究路径,应当暂时从纷繁复杂的视觉图像解析任务中抽身而出,首先就高层语义知识体对低层感觉信息的渗透模式做出一般性的理论研究,锻造出一个统一的技术平台。尔后,再把精力转向对这些技术平台的衍生品的开发上去。若不如此,对一个能够具有通用知觉能力的智能体的构建,恐怕亦将永远是一轮水中月,一朵镜中花。可惜的是,这一“从哲学到工程学”的工作进路,目前还没有在业界得到普遍的认同。

当然,本文的工作性质主要还是哲学的,而不是工程学的。在本文中,以维特根斯坦关于视知觉的丰富哲学评论为素材,笔者至少已经为这样的一个哲学论点做了辩护:

语义符号系统和原始感官信息之间固然有着自然的差异,但高层语义知识却有着在必要时干预低层感官运作的“特权”。换言之,并不存在着一个高层语义知识的推理规则无法渗入的感官信息层。由此看来,关于高层语义知识对低层感官信息的渗入方式,应当有一种统一的解决方案,而不能仅仅依靠对一系列彼此无关的技术路径所做的外在捆绑。

毫无疑问,关于如何将这样一种哲学洞见熔铸到实际的工程学构建中去,我们必须付出更为艰辛的努力。

(责任编辑:张琳)

Key words: reflective equilibrium; utilitarianism; intuitionism; foundationalism; naturalism.

• Exploring Late Wittgenstein's Philosophy from the Perspective of Reflective Equilibrium

CHEN Chang-shen

Abstract: When we explore how Wittgenstein deals with the “rule-following paradox” and analysis accuracy and ambiguity of concepts in *Philosophical Investigations*, and his “Copernican revolution” about the certainty of belief, and criticism of dogmatism and skepticism from the perspective of reflective equilibrium in *On Certainty*, it can be seen that there is a narrow reflective equilibrium between the rules of language-game (or behaviour) and specific language-games, and a broad reflective equilibrium between series of competing general principles and specific normal practice. These two kinds of reflective equilibrium are in accordance with John Rawls' three basic requirements and the five characteristics proposed by Follesdal. Besides, they also have some peculiar properties.

Key words: certainty; language; language -game; rule; rule -following; reflective equilibrium; Copernican revolution.

• Science and Enlightenment: Preimplantation Genetic Diagnosis and Discourse Ethics

Claude Debru

Abstract: Basic philosophical ideas regarding Enlightenment, such as Kant's discussion about the public use of reason in his writing “*What is Enlightenment?*”, were re-discussed by contemporary philosophers such as Michel Foucault and Jürgen Habermas, who both insisted on the political and ethical dimensions of the subject. Examples taken from the statements and discussions of the French national bioethical committee on issues like preimplantation genetic diagnosis are analysed in order to better understand the real process of argumentation and decision-making taking place in the committee in the spirit of Jürgen Habermas' discourse ethics.

Key words: Enlightenment; discourse ethics; preimplantation genetic diagnosis; bioethics

• What is Wittgenstein's Influence on the Investigation of Machine Vision?

XU Ying-jin

Abstract: “Machine Vision” (hereafter MV for short) can be the title of such technologies as imaging-based automatic inspection and robot guidance. It has evolved into a large family, each member of which has been developed for its own aim. In this sense, it is hard to identify a unified scientific basis of MV. In this paper, I shall argue that any MV approach must make assumptions. It is philosophers' task to reveal, test, criticise or justify them. Here, the controversial issue is: Does the information-processing principles in a vision module connect or separate from those in a central reasoning module? Under the influence of Wittgenstein's theory of perception, I shall argue for the continuity between vision and reasoning, and then propose a new approach for the investigation of MV.

Key words: Machine Vision; sense-data; visual perception; Wittgenstein; Gestalt principles.