

陈 静\*

(南京大学艺术学院)

## 人工智能时代的数字图像<sup>①</sup>与视觉知识生产

**【内容提要】**自 20 世纪以来,数字技术和数字媒介的应用使得大量图像被数字化,使得图像具有了新的属性、特征和施用范围,图像得以在数字环境中被调整、修改、编辑和再使用。同时,大量新的数字图像的出现,彻底改变了传统认知视野中的图像生成方式、呈现方式和其功能。由此,图像的内涵和外延都已经发生了变化。这种依托于新技术和新媒介的视觉语言系统及认知方式,被称为“视觉认识论”(visual epistemology)。本文正是以此为出发点,借鉴乔安娜·朱可的“图形发生论”(graphesis)理论,从数字媒介转型这个特定语境中去探讨数字语境中的图像、图像数字化以及图像语言元数据化等问题,试图从知识生产的角度去探讨图像作为一种视觉形式在新媒体环境下所具有的可能性。

**【关键词】**数字媒介;数字图像;视觉知识生产;数字化;元数据化

2018 年 7 月 18 日,美国谷歌(Google)发布了一款名为“猜画小歌”的微信小程序,短时间内风靡微信朋友圈,获得了广泛的关注。这款被称为“人工智能猜画”的小程序主要玩法是让神经网络在限定时间内识别用户绘制的一个日常物件的涂鸦,“旨在让用户了解、体验人工智能的乐趣”。其背后是一个“源自全世界最大、囊括超过 5000 万个手绘素描的数据群”。虽然这款带有游戏性质的小程序定位“在线教育”,但其教育的对象不仅仅是人工智能,也

---

\* 陈静,博士,南京大学艺术学院副教授。主要研究领域:文化与媒介研究,数字人文。本文为国家社会科学基金重大项目“西方美学经典及其在中国传播接受的比较文献学研究”(项目号 17ZDA021)的阶段性成果。

① 本文中的图像是指(image),graph 或者 graphic 在本文中被翻译为“图形”。

包括了用户。在程序的指令下,用户按照个人经验与常识绘制出花生、篱笆、自行车、棒棒糖、葡萄酒瓶或是彩虹的形状。当该图案被人工智能识别出来的时候,用户也确认了自己认知中的图示;当该图案无法被人工智能识别时,用户会在怀疑中自我调整,进一步去修改描绘。在这样的过程中,谁又在教育谁呢?谷歌公司通过这种方式获取了超过5000万个训练数据,不仅训练了谷歌神经网络系统,提升了算法,同时还说明了图像不仅仅是一种再现性的系统,其背后也是一种语言文化的影响与塑造。此后,以数据新闻著称的海外科技媒体 Quartz 和两位记者 Thu-Huong Ha 和 Nikhil Sonnad 则从人文视野出发,通过对来自谷歌开放数据库的、148个国家和地区的人所画的11.9万个不同的圆,以及每个玩家的手指(或鼠标)经过路线的坐标的分析,提出哪怕是绘制一个最一般而普遍的“圆”,其也并非是一种简单直觉的结果。事实上,在不同的国家、文化与语境环境中受教育的人们在绘制圆形也采取了不同的路径,比如在近5万个美国人所画的圆中有86%是逆时针画圆,800个日本人画的圆中有80%是顺时针画的。作者将其原因归因于不同语言的书写系统的书写方式,比如拉丁字母语言中成长起来的人习惯于逆时针,而书写像中文、日文的人则习惯于顺时针。而且,常用手也可能对塑造人们绘制抽象形状的方式产生影响。此外,作者还引用了自20世纪70年代以来的一些实验数据及结果表明,文化以及其中的书写方式会随着时间的推移越来越明显。更有意思的是,作者最后提出的问题是,当人们越来越习惯于通过敲打和打字来进行交流而忽略或者遗忘书写艺术的时候,是否会形成一种新的键盘引领的文化差异?<sup>①</sup>

这是人工智能时代图像背后的故事,也是乔安娜·朱可(Johanna Drucker)在论及数字时代的图像、视觉认知与知识生产时所谈到的:“视觉的再现方式已经过时了。我们现在已经知道是我们的感知承受程度与认知能力共同构成了对视觉世界的印象。我们所看到的世界是一个由我们认识能力所造就的世界”。<sup>②</sup>数字时代的到来已经改变了图像的生产、呈现和认知方式,同时也改变着具有视觉能力的人类主体认知能力。

关于图像的研究,长时间以来主要集中在三个领域,一种聚焦于由特定鉴赏能力和社会功能驱动的对艺术的风格(style)、图像学(iconography)以

<sup>①</sup> Thu-Huong Ha, Nikhil Sonnad, "How do you draw a circle? We analyzed 100,000 drawings to show how culture shapes our instincts", Quartz, 06-05-2017, <https://qz.com/994486/the-way-you-draw-circles-says-a-lot-about-you/>.

<sup>②</sup> Johanna Drucker, *Graphesis: Visual Forms of Knowledge Production*, Cambridge: Harvard University Press, 2014, p.19.

及其他形式特征问题；一种更多的关注于关于图画式图像的视觉文化问题；一种则关注于视觉传达（visual expression）的形式系统。这三种关于图像的研究尽管路径不同、方法各异，但却有着一个共同点，即在其领域中，图像并不以解释和分析知识生产为目的。即使在工程、建筑设计、地图学、经济学、统计分析和科学图示中，有着大量的通过图例来传达信息的可视化方式，但其目的依然是通过再现一种对可见或不可见的对象的视觉或者形态上的相似性再现或者表现来提供一种视觉形象。而对其的解释也一直从属与其所在学科的知识系统中，而没有被作为一个独立的领域所被认识。

信息及通信技术的发展以及数字媒介的渗透性应用，使得以往在高雅艺术、文化研究或者图形设计并不占据主要位置的、以解释和分析知识为主要目的的视觉知识系统变得更为突出和重要。一方面，数字技术和媒介的应用将已有的其他媒介的图像数字化，使得图像具有了新的属性、特征和施用范围，可以在数字环境中被调整、修改、编辑和再使用；另一方面，大量新的数字图像的出现，已经彻底改变了传统认知视野中的图像生成方式、呈现方式和其功能。由此，图像的内涵和外延都已经发生了变化，尤其是“图形”（graph 或者 graphic）开始溢出了其原有的意指边界，变得更多的流动和多义。像用户图形界面（UGI）设计、界面设计（interface）、生成图像（generate image）、交互设计（interactive design）、数据可视化（data visualization）、图像处理（image processing）和图像识别（image recognition）等视觉生产形式，不仅仅是传统认知中的对数学或者逻辑性知识的视觉再现或者传达，而是密切地与人们的视觉、感知、认知和文化传统习俗产生了深度关联。这种依托于新技术和新媒介的视觉语言系统及认知方式，被称为“视觉认识论”（visual epistemology）。本文正是以此为出发点，借鉴朱可的“图形发生论”（graphesis）理论，从数字媒介转型这个特定语境中去探讨数字语境中的图像、图像数字化以及图像语言元数据化等问题，试图从知识生产的角度去探讨图像作为一种视觉形式在新媒体环境下所具有的可能性。

## 一、图形认识论：从图像到图形

“图像”是一个内涵混杂且到处出现的概念。我们在英语中可以找到一系列关于图像的概念，比如 image、graph/graphic、picture、photo/photography、illustration、icon、chart 等，但在具体使用时，却往往并不加以细分。正如罗伯

特·米歇尔(Robert Mitchell)和雅克·卡利普(Jacques Khalip)所批评的:

“除了其显而易见的普遍存在性,‘图像’作为一种分析类型,依旧抗拒从理论上被理解。尽管文化批评、新媒体学者,以及社会学家对不同类型的图像进行了分类,描述了这些图像在文化中所起的作用,并说明了与图像相关的技术的特殊属性,但远还没有说清楚在这些研究中图像到底是什么,如何或者是否可以将图像与图片(pictures)、图标(icon)区别开来。”<sup>①</sup>这种概念上混杂性致使有关图像的讨论往往不是在同一范畴内展开,比如我们在讨论文学或者诗歌时候谈到的“非视觉的”图像性,这与艺术作品的图像性或者数字图像的图像性又都不同。

那么图像是否成了无法被言说的对象?事实上,尽管存在着如此复杂的分类系统,但从意义生产的内在机制上,可以根据意义的指向性关系将“图像”分为两种:一种是指向图像的内在逻辑。正如朗西埃所指出的,图像指向两种不同的事物,一种是生产了原初物的相似物的简单关系,一种是我们称为“艺术”的那种操作。前者不要求忠实复制而仅需简单有效地代表,后者则更多地涉及对这种相似性的修改(an alteration of resemblance)。<sup>②</sup>因此在内在逻辑上而言,图像代表的是一种关系,一种处于变动中的相像性,其有效性来源于原初物与再现物之间的关系的稳定性。如何使得这种相像性以及这种相似性关系得以可见(visible)、得以被认识,很大程度上取决于这种相像性、这种相似性关系究竟有多接近原初物。而其前提也是认知主体需要对原初物有所识别的前提下。

另一种则指向了图像的外在指涉。认知主体在看到图像的情况下,即使识别了图像,但可能依然无法理解图像,需要借助外在的语义系统或者文化语境。在这个意义上,图像是复杂社会文化中表意实践的产物,不能离开特定的社会语境、文本和主观建构。这种建构的方式与文化、社会和主体的建构方式有着密不可分的关系,因此在特定社会语境内的图像反过来也可以呈现出图像所在但又在其外的社会、文化、主体建构方式。如周宪所指出的,当观者从形象构成和媒介特性看待图像,实际上是将“形象(image)”细分为“图像”、“影像”和“景象”。而“表征”和“视觉性”与“形象”一起都成了视觉文化的一组具有“相似”的概念家族。因此,“形象”得以从单一媒介延展到文化

<sup>①</sup> Jacques Khalip, Robert Mitchell, eds., *Releasing the Image: From Literature to New Media*, Stanford University Press, 2011, p.3.

<sup>②</sup> Jacques Rancière, Gregory Elliott translated., *The Future of the Image*, London: Verso, 2009, pp.8—11.

领域，并且得以借助“表意实践”的方式被认识和分析。在这个意义上，周宪强调视觉性的社会意义，认为视觉性及“人的视觉的社会性，包括了‘视觉的社会建构’和‘社会的视觉建构’”而表征恰好是这两种建构导向的内在机制和外呈现方式。<sup>①</sup>

然而，在数字媒体时代，图像的两种逻辑在数字媒介所创造的环境中被打破了，而另一种具有新的特征的图形认识论由此产生——数字媒介中的图像更多的是以一种界面化图形方式的存在。

首先，数字媒介转型从内在逻辑上解构又同时重建了朗西埃意义上的图像相似性逻辑——图像以一种消解与唤醒同时存在的方式与其相似物建立起一种可识别的“简单关系”。此处引入博尔特（Jay David Bolter）和格鲁辛（Richard Grusin）的“再媒体化”（remediation）理论来在媒介语境中反思这种图像的相似性关系。该理论认为再媒体化过程具有去媒体性（immediacy）和超媒体性（hypermediacy）两种逻辑，从而使得图像在媒介转换过程中，特别是在新媒体环境中，承载了之前媒介的特性同时又具有新的媒介特征，从而使得观者不断地忘记媒介的在场、相信其与图像所再现对象一并在场的同时，又意识到媒介的存在。观看活动本身作为一种表征系统的使用和借用，不断地回到对原来媒介使用的感受中。以计算机和网络为代表的数字媒介将这种在前信息时代也存在的再媒体化过程推向了一种前所未有的极致。计算机程序的开发者努力消除自身在人与计算机的互动中的痕迹，力图给程序最大的自动性，从而达到去媒体的目的。比如，电脑操作系统“视窗”的发明“意味着计算机界面消隐入经验性的后台，软件所依据的相似、类比则走到了前台。如果绘图软件是‘直觉的’，那也是因为绘图箱有着一个在文化上相似的客体，从而具有了直觉。图形界面设计师在设计的时候不仅仅是依据了文化上相似的对象，而且是涉及了之前的媒体，比如绘画、输入和手写，创造了一套更为复杂的体系，在其中，视觉符号化的再现形式和表达形式相互作用”<sup>②</sup>。又比如综合了绘画、音乐、报纸、杂志、摄影及电视特征的万维网网页，模拟了手势动作的交互性游戏，文档处理软件等，无不是在新的虚拟空间中建立一种似曾相识却又完全不同的视觉语言和系统：人们在不断适应、调用这套新的系统的同时又不断地回想意识到这种“新”所带来的“不适应感”与新鲜感受。数字媒介在这种过程中扮演的不仅仅是一种文化研究意义上的表征系统的载体或是麦克卢汉所说的信息

① 周宪：《从形象看视觉文化》，《江海学刊》2014年第4期，第177—185页。

② Jay David Bolter and Richard Grusin, *Remediation: Understanding New Media*, the MIT Press, Cambridge, Massachusetts, London, England, pp. 272—273.

(语符)的组织方式,更是基于主体认知的方式对视觉系统进行再生产,从而产生一种基于数字媒介的视觉知识机制和系统。

其次,数字媒介从生成方式、传播方式、呈现方式上实现了数字化、图形化和界面化,形成了一个自在自为的闭环,改变了作为图像意指外在指涉参照物的文化—社会意指实践系统。从生成方式而言,在新媒体语境中图像一般可以认为有两种:一种是通过数字化(digitization)的方式将在非数字化环境内生产出来的图像转为数字图像,另一种是数字生成的(born-digital),是在数字化环境中直接生成的图像。但不论是文本、图像、图形还是声音文件,都是用计算机代码写的一套有效的数学指令,即1和0的字符串,也就是作为电子脉冲传输中两种不同的电压。因此,数字媒介实为“将离散单元创造成一个系统的中间代码。数字代码将所有东西都分解成一致的可交换的比特(或字节)。而数字代码可以解释聚合媒体和创造新媒体中出现的大多数激进可能”。<sup>①</sup>数字媒体以一种系统的中间代码,即二进制代码来表征对象。在这个表征系统中,既不存在表象上的差异性,因为所有的代码都是1和0的字符串,也不存在“物质性”的差异,因为所有的二进制代码“还原”成了电路或者电子脉冲传输中的两种不同的电压。因为这种电子化的过程是无法被直接看到或者感受到的,所以该表征系统对于其所生成的媒介(比如网络)来说是“虚拟的”。这种虚拟性恰恰造成了观者对于媒介差异性的无感。弗雷德里奇 A. 基德勒(Friedrich A. Kittler)将这一过程描述为——信息和信息渠道的普遍数字化抹除了单个媒体间的差异。音响、图像、声音与文本都变成了界面效果。<sup>②</sup>从传播方式而言,数字化或者数字生成的图像都可以通过数字化渠道传播的,特别是数字生成的图像大部分只能在数字化渠道中生存,无法转为其他媒介形式。那些可以转为其他媒介形式的,比如数字编辑排版设计的书最后可以通过印刷方式成为纸质的书籍,或者,数字化拍摄的照片可以打印在不同材质上成为基于不同媒介的实体照片,也是因为这些图像可以“界面”的方式转换为另一种媒介的形式;从呈现方式而言,原则上,所有外在于人眼睛及大脑生理功能的图像都要通过一种物理性“界面”呈现,比如书籍是通过书页,油画是通过画布,壁画是通过墙壁,电影中的活动影像是通过胶片、屏幕及银幕等。数字化或数字生成图像除了转为传统媒介呈现方式外,更主要地通过摄像机的小屏幕、计算机终端、手机终端、显示器等各种终端装置来呈现。可以简单地总结为:数

① Andrew Dewdney and Peter Ride ed., *The New Media Handbook*, Routledge, 2006, p.228.

② Friedrich A. Kittler, “Gramophone, Film, Typewriter”, *October*, 1987(41), p.101.

字时代图像的生产方式单一了，传播方式反向逆转了，数字图像成了原图，而界面的多元化导致阅读方式的多样性。

## 二、图形语言：图文关系重构

尽管如前所述，无论是文本还是图像，在数字媒介中都成了一种界面原图意义上的图形，但我们无法改变不同媒介中数据流有所区别的事实：即使是在新媒体时代，无论是数字化的或者是数字生成的图像依然有着不同于文字文本、语音或音响等的特性。

印刷时代也有纸质媒介中对本雅明所称为的“视觉性材料”(visual material)的借用——在印刷文本中也有图像等视觉元素，例如书籍中的插图、装帧、封面设计，等等。特别是在中世纪手抄本中，图像与文字发挥着几乎同样重要的作用。但在印刷技术成熟之后，视觉元素在书籍中的作用就逐渐降低，成为文字补充的次一级存在。纯粹的文字所具有的图像性效果，仅仅在于书籍的空间设置上，包括页边、顶部和侧边、空页、标题页和封底的空间，尾注和脚注、目录和附录、页面和装订以及字之间、例子周围，单独的字母、字距、缩进、空格、字形之间的空间。图像成了文字背后的潜在层面。尽管如沃尔特·翁所说的，在从口语到书写的过程也是一个听觉主导转向视觉主导的过程。但是在这里视觉主导更重要的是指对于印刷文本的阅读在从口语走向默读之后，观看比聆听更为重要。<sup>①</sup>在印刷文本中，其所具有的图像是通过文字的印象来实现的。W.J.T.米歇尔将之称为“文本图像”，“同样，我们在语言表达中发现的‘图像’，无论是形式的还是语义的，都不被直义地理解为图像或视觉景观。它们只是相像于真正的画或视觉形象——被双重稀释的‘形象的形象’或我在别处所说的‘超图像’……我认为我们会遇到看和说、画和印刷汇聚于叫作‘文字’的媒介上来”。<sup>②</sup>文字替代了视觉的表征功能，而称为一种语义关系。我们在文学作品中看到“活生生”的文字描述都是为了能够激发个体的想象，从而能够在大脑中幻想出一个图像化的对象。这种图式化的文字与图像的关系在19世纪经历了一次变革，“其中，诗歌抛弃了与绘画的联合，在音乐中找到了新的类比，这已是人们耳熟能详的了”。通过文字“想象”在浪漫

① 参见沃尔特·翁，《口语文化与书面文化——语词的技术化》，北京大学出版社2008年版。

② [美]米歇尔：《图像理论》，北京大学出版社2006年版，第133页。

主义诗人那里被赋予了超越视觉化的力量。“可视语言”在此时期达到了一个巅峰,文字与图像之间的裂痕也进一步加深。

在数字媒体时代到来之后,这种文字与图像的关系发生了重大的改变。图形成了与图像、文字一样的文本,并真正的从文字的主宰中解放出来,不再成为文字背后的隐藏层面,需要借助文字去理解和想象,而是恢复了图像自身所具有的视觉性力量。自视窗与超文本时代起,图形就已经以一种视觉隐喻的方式打造了一个图像化的书写空间。在这个空间中,甚至文字也成为了图像。或许有人会说,即使是在电子作品中,对于图像的解读依然要依赖于语言,但事实上,随着对于代码和语言的关系的进一步认识和探索。人们开始重新认识“语言”,这里的语言不仅仅是人类的自然语言,还包括了图形语言(the language of graphics),甚至机器只读的代码,这就从根本上对于图像和文字的关系进行了再建构,图形界面、机器代码与程序编程作为一种语言也参与到了文本的意义之中。比如图像标注系统,一般包括人工标注和机器自动标注两大类。人工标注主要借助领域专家、元数据框架和领域主题词表进行。近年来,大众标注(Folksonomy)逐渐兴起,如Flickr网站,它借助普通大众和自由的tag标签,而非严格的主题词实现海量图像的标注与组织。尽管大众标注使用的词汇并不规范,但是它解决了专家标注速度慢和标注词汇自由度不足的缺陷。

在这里,可以举一个《威廉姆·布莱克档案库》的例子来说明图形语言是如何在图像数字化过程中通过“再媒体化”方式在图形界面上实现对图像与文本的关系进行重构的。

威廉姆·布莱克作为18世纪末到19世纪初的一位重要的印刷商、雕刻家、画师和诗人,他倾注平生的精力,努力将语言综合在诗歌和绘画的合成艺术中。他的作品可以说是在那个时代独特手稿和普遍传播文本、诗歌与绘画、视觉与文字的关系的潜能所能作出的最具创造性的贡献。在布莱克的插图中,他的文字与图像融为一体,在插图中字母是经过专门的设计,他用的字母常常生长出枝蔓,只能根据图像来解释。同时,他插图中的所有图像,小到一朵花边或者页边的漩涡大到整体主题和形象的选择都具有体现着插图中的“文字”所代表的意义。布莱克在纸质印刷媒介时代重构了文本与图像的关系,改变了图像作为文本的“次文本”的地位,并利用了不同媒介的特性表达了他的文本观——在一个整体性的文本中融合言语与文字、图像与文字、听觉与视觉以及想象和理性等相互对立的一系列关系,体现一种“超文本”的特征。但是,值得注意的是,布莱克对于图像文本与文字文本的融合是通过一种形而上学的象

征主义精神达到的，这就意味着，对图像文本和文字文本的解读一样要依赖于解释语言。应该说，布莱克的文本是以一种颠覆性的叛逆姿态对印刷术及书籍文化进行了最大可能的试验。布莱克作品所体现出的丰富性也使得在进入数字时代之后，将布莱克的作品电子化成了学者们最早关注的项目之一。

威廉姆·布莱克档案(The William Blake Archive, <http://www.blakearchive.org/>)就是这样一个在线的数字档案项目，由三个当代最突出布莱克研究方面的学者和编辑设计的。该档案本着很高的学术规范要求，与许多主要的博物馆、图书馆和私人收藏家合作，利用最先进的技术收录了大量布莱克的插图书和其他一些商业或者创作作品的高质量复制品。同时，该档案还利用超媒体系统才能提供的对图像和文本的组织、内在链接以及搜索方式等学术手段，在尽可能符合布莱克视觉和文学规范的前提下第一次为学者和学生们访问布莱克作品的机会。从技术上来说，这不仅是一个杂糅的、多合一版本的、分类的数据库，还是一套学术工具，这套工具能够充分利用新的信息技术所具有的优势。就该档案目前提供的布莱克作品和文本的数量、质量及其编辑、设计的水平而言，说威廉姆·布莱克档案树立了一个文学网站的黄金标准是毫不为过的。

作为一个充分利用了信息技术的多媒体网站，该档案的编辑们并不是简单地将布莱克的作品电子化，而是采取了一定的编辑方式，设定了一些内部链接，以建构出布莱克图像和文本之间、文本与文本之间的一种关联性。在这个可以称作凯瑟琳·海勒丝(N. Katharine Halves)意义上的“翻译媒体”的过程中，这些编辑声明他们是将“作品”视为一种独特的物理对象，即注意到了文本的物理特点——页面大小、字母、衬页、行距等“书目代码”(Bibliographic Codes)，将之视为与语言符号一样的意义元素。而在对文本的电子化处理上，编辑们聪明地使用了计算机的仿像功能从而尽可能地让屏幕显示看起来像印刷书一样。并且还向用户提供了 Calibration Applet 这样的 Java 小程序来设置屏幕以再现原始书页的维度。由于档案的编辑们小心的坚持物质性对意义的潜在影响中产生的细微差别，因此他们需要克服很多麻烦去搜集不同的作品或者同一作品的现有的几种副本。然而这些副本借用技术在屏幕上的视觉呈现在其物质性上与印刷之间的差异远远大于不同的印刷副本之间的差异。从内在技术层面而言，电脑对印刷文献的精确模仿恰恰就是因为它在其架构和功能上是完全的不同于印刷。视觉精准度的模仿在将布莱克从单一文本版本中解救出来，而后者压抑了他作品的重要的视觉维度。

从这个意义上而言，该档案是一个多媒体的威廉姆·布莱克百科全书。如果说，布莱克自身在努力地通过一系列努力在他的纸质文本中重构着图像与文

本、言语与文字、声音与听觉、手稿与印刷等关系。那么布莱克档案则将他的这种颠覆性的“泛文本性”真正推进到了超文本结构的图形化:(1)文字与言语之间的裂痕在数字化的空间中被进一步的弥合。手稿和印刷都成了一种“源代码”,在多媒体空间中被重新编码,形成新的对照关系,并且在传播上达到了印刷书所达不到的广泛程度;(2)如果说布莱克是以象征性的方式来表达对一系列浪漫主义的观念以及线性叙事的权威性的反抗,那么在档案中,这种象征性的方式被数字化后现实化了,人们以与身体相关的沉浸性感受“游历”;(3)图像不再是文本的隐形层面或者“次文本”——图像和图像之间、图像和文本之间、文本和文本之间、图形与图文之间形成了一个离散的网络,文片间呈现出互文性的关系,彼此补充、互为说明;(4)原有的声音对印刷、古代文本对现代文本、想象对理性的对立性被消解了。超文本去中心化的结果就是离散化的文本存在;(5)从阅读来说,布莱克档案提供了印刷文本所不具有的参与性。读者可以根据导航栏和链接等方式自己选择浏览路径。

### 三、视觉知识生产:图像的元数据化

尽管布莱克档案体现了数字图像与文本的高度图形化结合,但不可否认在这个过程中,图像的意义生产方式无法完全脱离传统的意义生产方式,依然要高度依赖语音中心主义线性逻辑。但我们已经可以看到,图形化的图像更多地向数字媒介方式打开,开启了新的视觉知识生产的可能性。而这一点在图像的元数据中将体现地更为淋漓尽致。

在《元数据,我的爱》(*Metadata, Mon Amour*),列维·曼诺维奇认为“元数据化图像”为“界面现实”(interface reality)和新型的人类经验提供了一种新的范式,其可延伸到以下四个方面:(1)创造新的图像描述和分类系统;(2)为图片收藏创造了新的界面;(3)创造了新图像类型,突破我们都已经熟知的“静态照片”或者“数字视频”等类型;(4)使得网络上那些达到了“超人类”规模的视觉数据不再是一个问题而成了创造性的机会<sup>①</sup>。在这个议题的讨论中,曼诺维奇教授的两个观点/立场非常值得注意,他强调了数字技术(包括数字相机、处理器、扫描仪及各种软件)对于生产数字图片元数据的重要

<sup>①</sup> Lev Manovich, *Metadata, Mon Amour*, <http://manovich.net/index.php/projects/metadata-mon-amour>.

性，同时他还提到数字图片的规模使得我们必须要考虑从“抽样”到“完全记录”的转变，这是因为传统的叙事艺术和现代媒体技术（电视、电影等）事实上都是基于抽样现实（sampling reality），也就是说，它们仅仅再现/记录了人类经验的细小碎片，那么数字记录和存储技术则极大地扩展了可以被再现和记录的对象的范围。

非常有意思的是，曼诺维奇还进行了相关图像数据库实践。早在2009年，他就在加州大学圣地亚哥分校的软件研究创新中心（Software Studies Initiative at University of California, San Diego, UCSD）和加州电信和信息科技学院（加州电信及信息科技学院，Calit2）开展了一个以日本卡通漫画为研究对象的数字项目（<http://lab.softwarestudies.com/2010/11/one-million-manga-pages.html>）。

他的团队从日本漫画迷网站 OneManga.com 上下载了 883 个系列的漫画，共是 1,074,790 个漫画页面。每个页面都是 JPEG 格式，850×1150 像素。他们在美国国家能源研究科学计算中心的超级计算机上运行了他们定制的图像处理软件测量了每页上的视觉特征（明暗、细节呈现，质感），然后运用了他们称为“文化分析”（Cultural Analytics）的方法，对图像已有的元数据，比如出版日期、作者姓名，和通过后期标注或者编码的新的元数据进行数据化再视觉化的呈现，用图示来表示对这些漫画图像的研究结果。

曼诺维奇的理论和实践的意义在于：首先，图像对于叙事艺术一直是附属性质的，其重要性在于“图例”和对现实/人类经验的反映和再现，这点在新的数字图像时代已经被扭转了，图像应该具有不同于语音中心主义的“语法逻辑”和语言，也就是新的图像描述和元数据模式与分类；其次，数字技术的发展不仅促使大量数字图片的出现，而且还使得数字图形成重要的人类经验的记录方式，数字技术本身就参与到了图像元数据的建构中；第三，当考虑到当代所有的文本都可以成为数字（digit），那么图像与图形化的界面事实上成了信息视觉化最重要的方式，也成了人们与信息及信息机器交流的最重要的途径。

由此，人与计算机的交互作用已经改变了“人”作为意义生产的唯一主体。正如基特勒指出，意义并不是主体所赋予技术的，不是内在或者先于技术存在的，而是由技术所引导的，或者因为技术而可能的<sup>①</sup>。走得更远一点，即媒介的本体论也就成了地址、指令和数据的三位一体，是在“硅固体物理学和冯

---

<sup>①</sup> 参见 Friedrich Adolf Kittler, *Gramophone, Film, Typewriter* (Stanford University Press, 1999)。在这点上，我认为基特勒是追随了本雅明的脚步。本雅明观察到了每个历史的时期，某种特定的艺术形式只有通过一种技术标准的改变才能实现其全部的影响力。

诺依曼体系结构双重条件之下的媒介本体论”。<sup>①</sup>如果我们认为基特勒是对的，那么图像的意义生产方式最终走向的是由图像作为知识生产的自发性、自觉性和自律性。

---

<sup>①</sup> 弗里德里希·基特勒：《走向一种媒介本体论》，胡兰菊译，《江西社会科学》2010年第4期，第255页。

background of Geospatial Turn in the hope of correcting Western thought theory terminology description used in academic circles theoretically, making it more accurate and comprehensive. Simultaneously, it has important theoretical significance that the nomenclature correction of the academic terminology description of western academic thought trend in recent years.

**Key Words:**

Geospatial Turn; Geospatial Criticism; Western Academic Thought Trend in Recent Years; Human Geography; Critical Human Geography.

## **Digital Image and Visual Knowledge Production in the AI Era**

■ Chen Jing

(Nanjing university)

**Abstract:**

Innumerable images have been digitized or born digitally with the application of the digital technology and media since the late last century. Being gained the new attributes, features and usages, digital images could be adjusted, edited, revised and reused in the digital environment, which not only challenged and changed the traditional way of epistemological understanding of the production, representation and function of images, but also produced a new system of visual language and a approach of recognition of images based on the new technology and media, the “visual epistemology” . This paper takes it as a starting point to discuss the fundamental issues of definition, digitalization and metadating of images during the transformation of media to explore the potential of images as a kind of visual form of knowledge production in the new media environment.

**Key Words:**

Digital Media; digital image; Visual Knowledge Production; Digitization; Metadating