

• 研究前沿(Regular Articles) •

基于双眼视差的立体视觉去掩蔽效应*

陈明立¹ 张畅芯¹ 杨少娟¹ 毛利华^{1,3} 田永鸿^{2,3}
黄铁军^{2,3} 吴玺宏^{2,3} 高文^{2,3} 李量^{1,3}

(¹北京大学心理学系; ²北京大学信息科学技术学院; ³教育部机器感知与智能重点实验室, 北京 100871)

摘要 基于双眼视差的立体视觉不改变目标与掩蔽刺激之间的信噪比, 但能使不同的刺激被知觉在不同的深度位置上以降低目标信号所受到的掩蔽作用。本综述在总结前人双眼去掩蔽研究的基础上, 强调深度维度上的视觉注意在这种主观空间分离去掩蔽过程中所起的重要作用, 并介绍了双眼去掩蔽在现代科技领域中的典型性应用。最后, 结合听觉主观空间分离去掩蔽的研究进展, 本综述认为主观空间分离去掩蔽是大脑处理复杂刺激场景的一个基本功能。

关键词 双眼去掩蔽; 视差; 立体视觉; 随机点立体视图; 主观空间分离

分类号 B842

1 前言

Changizi 和 Shimojo 在 2008 年的 *Journal of Theoretical Biology* 期刊上发表了一篇理论性文章, 对动物双眼在头颅上位置分布的生态学意义作了理论分析。他们认为, 对于生活在没有视觉遮挡环境中的动物, 眼睛分布在头颅的两侧可以获得全景视觉(包括可以看到身后的刺激)并且不需要双眼之间的会聚(convergence)。然而, 生活在有视觉遮挡的环境(如有叶子环境, leafy environment)中的动物(包括人), 两眼分布在头颅的前侧可以减少遮挡物(如树叶)的遮挡作用。并且 Changizi 和 Shimojo 将由双眼之间位置的分离所形成的能“透过”遮挡物而看到遮挡物后面物体的能力称为“X-光”视觉(X-ray vision)。他们进一步认为, 这种“透视”能力与瞳孔的间距和遮挡物的尺寸之间的比例有关, 即在树丛中如果瞳孔间距大于树叶的尺寸, 那么前视的动物可以透过树叶而观察到其

后面的物体。尽管 Changizi 和 Shimojo 的 X-光理论(X-ray hypothesis)强调了双眼视觉的去掩蔽这一重要功能(有关视觉掩蔽的基本概念见综述: Breitmeyer, 1980; 陈斌, 高闯, 王建中, 2009), 但我们认为这个理论在对双眼视觉去掩蔽机制的认识上还有很大的局限性。

实际上, 具有前视特征的灵长类动物(包括人)以及其他许多动物(尤其是大多数夜间活动的捕食动物)的双眼感觉信号的整合不但形成了左右眼视野之间的重叠(或部分重叠), 而且也使得物体在左右眼视野中的视网膜成像之间会因为两眼之间的位置分离而存在一些微小的差异, 即双眼视差(binocular disparity)。视觉中枢系统可以从这种二维的视网膜双眼视差信号中提取产生立体视觉的信息(Barlow, Blakemore, & Pettigrew, 1967)。由于立体视觉的出现, 具有不同视差的视觉客体(如随机散点)被知觉在不同的深度面上, 因而观察者可以在感受到较近的主观深度面的同时也感受到了较远的主观深度面, 继而形成了所谓的立体透明性(stereoscopic transparency)的主观感受(Akerstrom & Todd, 1988; McKee & Verghese, 2002; Tsirlin, Allison, & Wilcox, 2008; Tsirlin, Wilcox, Allison, 2010; Wallace & Mamassian, 2004)。

收稿日期: 2012-02-05

* 科技部“973”项目(2009CB320901)、科技部国际科技合作项目(2010DFA31520)以及国家自然科学基金面上项目(31070897)资助。

通讯作者: 李量, E-mail: liangli@pku.edu.cn

根据 Tsirlin 等人(2008)的研究,在没有其它单眼线索的条件下,人类被试单靠双眼视差线索最多可以产生 6 个立体深度面。

立体视觉在人和动物对环境的适应过程起着重要的作用,因为立体视像的形成不但可以使观察者能更加准确地判定物体的深度距离(Cartmill, 1974),而且使对前肢的运动(如捕猎和抓取动作)的控制更加精准(Collins, 1921)。然而,随着对人类以及动物双眼立体视觉研究的深入,已有大量的证据表明双眼视觉的功能不仅仅在于准确定位和运动控制,而更重要的是,双眼视觉增加了观察者在复杂环境中消除伪装并发现目标客体的机会(Heesy, 2009; Julesz, 1964; Nakayama, Shimojo, & Silverman, 1989; O'Toole & Walker, 1997; Wardle, Cass, Brooks, & Alais, 2010)。

针对双眼立体视觉的去掩蔽功能,本综述在总结前人相关研究工作的基础上,强调双眼视觉去掩蔽的心理机制的重要作用,即在深度维度上所形成的前后主观空间分离以及注意在深度维度上的分配这两方面心理过程是双眼立体视觉去掩蔽的两个重要基础。

2 基于双眼视差的立体视觉去掩蔽

2.1 什么是“双眼去掩蔽”(binocular unmasking)

早在上世纪 60 年代, Julesz (1964) 等人就已报道了基于双眼视差的立体视觉去掩蔽效应的存在。Julesz 等人使用随机点立体视图(random-dot stereogram, RDS)为实验刺激,发现在没有双眼视差或者单眼观看的情况下,观察者无法觉察到散点背景中的一个方形。然而,如果水平视差线索被引入以使得一个散点方形的视差与其周围背景散点的视差有所不同,则在立体镜下观察者可以很容易地看到一个由散点组成的方块或浮于散点背景的前方或凹陷于散点背景的后方。

进一步的研究表明,双眼视差在引发观察者立体感的同时也显著地增强了对被掩蔽目标信号的觉察和识别。即当目标信号受到干扰刺激的影响时,如果刺激中含有能使目标视像和干扰刺激视像前后分离的双眼立体线索,那么对目标信号的觉察/辨认/搜索成绩要显著地高于不包含这些线索时的成绩(Moraglia & Schneider, 1990, 1992; O'Toole & Walker, 1997; Otto, Bach, & Kommerell,

2010; Speranza, Moraglia, & Schneider, 1995, 2001; Wardle et al., 2010)。例如, Moraglia 和 Schneider (1990)让被试觉察在高斯噪音图像掩蔽下的一个 Gabor 信号。在一种实验条件下噪音和信号之间存在双眼视差的差别(即噪音和信号被知觉为处于不同深度平面上),而在另外一种实验条件下噪音和信号不存在这种差别(即噪音和信号被知觉为处于同一深度平面上)。结果表明,当信号与噪音之间的双眼视差不同时被试觉察信号的阈限显著地低于信号与噪音处于同一视差水平下的阈限。因为两种实验条件之间的唯一不同是相对视差的差别,因此 Moraglia 和 Schneider (1990)认为噪音与信号之间的双眼视差的不同造成了探测成绩的不同。Moraglia 和 Schneider (1992)称这种去掩蔽现象为双眼去掩蔽(binocular unmasking)。

这种双眼去掩蔽效应不仅存在于人类视觉系统中,一些动物被证实也具有这种能力。例如, van der Willigen (2011)通过一系列行为实验证实了具有高度双眼视觉能力的猫头鹰也可以利用双眼视差线索从立体散点图中识别出三维立体目标信号,即从嘈杂背景中探测目标。

2.2 双眼去掩蔽效应中刺激线索的量化范围以及去掩蔽的程度

已有实验证据表明,视觉刺激信号的空间频率影响双眼去掩蔽的程度(Henning & Hertz, 1973; Moraglia & Schneider, 1990; Otto et al., 2010; Schneider, Moraglia, & Jepson, 1989)。例如, Henning 和 Hertz (1973)发现当目标光栅的空间频率为 0.24、0.6 或 2.4 cpd (cycles per degree)时,可获得去掩蔽量在 5~15 dB 之间(即视差的引入而降低觉察阈限的程度,以分贝(dB)为单位)。而光栅的空间频率进一步增大时,去掩蔽量逐渐降低,直至当空间频率增大到 6 cpd 时去掩蔽现象消失。Schneider 等人(1989)也研究了在视差恒定的条件下目标信号的空间频率对双眼去掩蔽的影响,并发现 Gabor 信号的空间频率在 1.1、2.2 或 4.4 cpd 情况下可以获得 6~12 dB 的去掩蔽量。因此,双眼去掩蔽的程度是很显著的。但去掩蔽效应主要针对空间频率不是很高的成分,对高频的精细结构成分作用不显著。

Morgalia 和 Schneider (1990)进一步研究了掩蔽刺激的不同视差对去掩蔽量的影响。他们的结果表明,当视差为零的目标刺激的空间频率为 2.2

cpd 时, 将噪音的视差设为 13.52、40.56 或 67.60 弧分(arcmin)时可以获得 6-12 dB 的去掩蔽量。

Wardle 等人(2010)用无色彩自然实物图像作为掩蔽刺激并以该掩蔽刺激中央部分的一个相位随机区域为目标刺激来研究双眼视差的去掩蔽作用。在其中的一个实验中(Experiment 1A), 掩蔽刺激的双眼视差为零, 而目标刺激的双眼视差为零、交叉(-6 弧分)或不交叉(6 弧分) (有关交叉和不交叉双眼视差的概念见下一节)并采用自由观看(free viewing)的刺激呈现方式, 刺激的呈现时间为 1 秒。这个实验的结果表明, 当目标刺激与掩蔽刺激的双眼视差不一致时, 对目标的觉察阈限要显著地低于当目标刺激与掩蔽刺激的双眼视差都为零时的阈限。具体来说, 在目标视像远于掩蔽视像和近于掩蔽视像的两种条件下的去掩蔽程度可分别达到 7.7 dB 和 9.0 dB。

此外, 在另一个实验中(Experiment 1B), Wardle 等人除了将目标刺激的双眼视差设置为零以及将掩蔽刺激的双眼视差设置为零、交叉或不交叉之外, 也对双眼的会聚活动进行了控制。他们将刺激呈现时间缩短为 200 ms, 并且只在双眼融合形成时(当分别呈现在左右视野的竖直游标线重叠后)再由被试自己按键来启动实验刺激。在这种实验条件下, 目标视像远于掩蔽视像和近于掩蔽视像所造成去掩蔽程度分别为 4.4 dB 和 5.9 dB。与上一个实验结果相比, 这一实验的结果反映了刺激呈现时间也会对去掩蔽的程度有所影响。

2.3 交叉性视差与不交叉性视差信号的不等同效应

双眼视差的方向性决定了所知觉到的立体视像是在注视点的前面(靠近观察者)还是后面(远离观察者)。具体来说, 当物体在左右视网膜上的成像相应于注视点有向外侧位移而使观察者知觉到的视像位于注视点的前方(更近于观察者)时, 相应的双眼视差模式为交叉性视差(crossed disparity); 反之, 当物体在左右视网膜上的成像相应于注视点有向内侧位移而使观察者知觉到的视像位于注视点的后方(更远于观察者)时, 相应的双眼视差模式为不交叉视差(uncrossed disparity)。

尽管一些研究结果显示在交叉性视差与不交叉视差条件下对信号的觉察阈限没有显著差异

(Morgalia & Schneider; 1990; Uttal, Fitzgerald, & Eskin, 1975; Wardle et al, 2010), 但也有研究结果显示在交叉性视差条件下对信号的觉察时间阈限要显著地低于不交叉视差条件下的觉察时间阈限(Manning, Finlay, Neill, & Frost, 1987)。因此, 一个有意思的问题是, 在这两种类型的视差条件下的行为成绩是否有所不同? 针对这个问题, Patterson 等人(1995)试图通过系列化的实验来得到明确的答案。他们的简单觉察任务结果表明, 在交叉性视差条件下对信号的觉察时间阈限与不交叉视差条件下没有显著性差异。然而, 在交叉性视差的条件下的深度知觉成绩则要显著地高于在不交叉性视差条件下的。研究者在迫选任务中让被试判断散点图中两个并排的立体方块的相对深度, 结果表明在交叉视差条件下辨别相对深度的时间阈限要显著地低于在不交叉视差条件下的, 而且在交叉视差条件下深度感更加真实。因此 Patterson 等人(1995)认为, 当加工水平涉及到深度计算和辨别任务时, 视觉系统对交叉视差和不交叉视差的处理机制在时间整合特性上存在差异。Tsirlin 等人(2012)用两组具有不同视差的随机散点来形成两层立体散点视像。研究结果表明, 当散点的总数不变时, 当在距离观察者较近的主观深度层面的散点数目比例越小时(即在距离观察者较远的主观深度层面的散点数目比例越大时), 主观立体分层就越困难。进而从另一个方面反映了双眼立体视觉加工在深度维度上的梯度。我们对 Tsirlin 等人用来解释这种现象的信号强度(signal strength)假设持怀疑态度。结合 Patterson 等人(1995)的研究结果, 我们的假设是: 在立体视觉形成过程中, 要根据靠近观察者的主观深度平面为参考面来形成不同深度面的主观感受。这样, 当在靠近观察者的主观深度面上的散点稀少时, 这个参考面就被弱化, 进而导致了深度感的弱化。

有关双眼视差的脑成像研究也表明, 视觉皮层对交叉性视差和不交叉性视差的反应是不对称的, 例如交叉性视差刺激在 V1\3\4 等脑区引发的反应较大(Adams & Zeki, 2001; Cottureau, McKee, Ales, & Norcia, 2011, 2012; Preston, Li, Kourtzi, & Welchman, 2008)。

需要指出的是, 交叉性视差与不交叉性视差实际上反映了视觉客体与注视参考面之间的相对

深度位置。在 Devisme 等人(2008)的由视差所引起的曲面变形的实验中可以体现出这种深度位置的相对性的意义。在他们的实验中,给被试呈现散点立体视图。在每一个试次中,散点图的中心区域始终保持为零视差,而在外周区域则引入非零视差值,并且从中心区域向外周区域随着离心距离的增大而使视差值有线性的增加。Devisme 等人记录被试知觉到散点图由平面变成曲面时的视差值,即形变探测阈限(the threshold of surface-deformation detection)。外周区域的视差可分为两种:交叉和不交叉。在外周区域为交叉视差的条件下,曲面被感觉为凹离观察者;而在外周区域为不交叉视差的条件下,曲面被感觉到凸向观察者。实验结果表明,当视差从中间区域的零值而逐渐向外周的不交叉方向增大时,被试的形变探测阈限显著地小于外周交叉视差条件下的形变探测阈限。这表明,相对于形成凹离观察者视像的视差,视觉系统对形成凸出视像的视差更加敏感。因此,交叉性视差与不交叉性视差之间的差异是相对深度位置差异的反映。

2.4 双眼去掩蔽过程中目标和掩蔽刺激之间相对主观位置的效应

使用立体镜来形成的基于双眼视差的立体视觉中,目标视像与掩蔽视像之间在深度维度上的(主观)分离并不影响单眼刺激强度的信噪比。因此,由这种主观空间分离所产生的去掩蔽效果一定发生在中枢(而非外周)加工阶段(Fox & Patterson, 1981)。许多研究工作表明,相对于所知觉到的掩蔽刺激所在平面,当目标的视像出现在前平面上(更靠近观察者)时,对目标的觉察成绩比在目标的视像出现在后平面上(更远离观察者)时的要显著地高,而且所需要的反应时间也显著地少(Fox & Patterson, 1981; Lehmkuhle & Fox, 1980; Nakayama & Silverman, 1986; O'Toole & Walker, 1997; Patterson et al., 1995)。例如,Lehmkuhle 和 Fox (1980)以及 Fox 和 Patterson (1981)用短暂呈现的兰道环视标 C (Landolt C)为目标刺激(被试需要辨认该刺激开口的指向)并用圈住该目标刺激的圆环为掩蔽刺激,发现由双眼视差所形成的目标刺激与掩蔽刺激在深度上的主观空间分离可以显著地减少对目标刺激的掩蔽。而有趣的是,当目标刺激的主观位置较掩蔽刺激更近于观察者时,被试的 C 环缺口指向的辨认成绩高于当目标

刺激与掩蔽刺激被知觉在同一深度上的,并且随着掩蔽圆环与目标在深度上的主观分离加大,辨认成绩也进一步提高(即掩蔽程度进一步下降)。然而,当目标刺激的主观位置较掩蔽刺激更远离观察者时,被试的 C 环缺口指向的辨认成绩要低于当目标刺激与掩蔽刺激被知觉在同一深度上的。这些研究者将这种在相对深度上不对称性的掩蔽效果称为“前方效应”(front effect),并假设造成这种效应的原因是因为靠近观察者的刺激比远离观察者的刺激要占用更多的视知觉加工资源。由此可见,若观察者的觉察任务不需要太多的加工资源,则不出现目标与掩蔽刺激前后相对主观位置的不同所引起的觉察成绩的显著性差异。只有在任务难度较大的条件下这种相对主观位置的效应才出现。

视差的不同造成了目标刺激客体与掩蔽刺激客体在深度上的相对远近的不同,这种效应也可以由掩蔽场景下的搜索任务中反映出来。例如, O'Toole 和 Walker (1997)考察了在视觉搜索任务中交叉视差和不交叉视差对反应时间的影响。他们采用 RDS 作为实验材料,并使散点图中包含了若干个由视差所形成的立体方块。这些方块或者都在同一深度平面上(没有目标方块),或者有一个方块(目标方块)处于不同于其他方块(掩蔽方块)的深度平面上。在刺激呈现时间不受限制的条件下,被试搜索这些立体方块并判断是否有目标方块的存在。实验中通过改变干扰方块与目标方块的视差方向(包括干扰方块与目标方块的视差分别为交叉和不交叉,以及不交叉和交叉两种条件)来调节目标方块和干扰方块的相对深度位置。结果显示,随着干扰方块数量的增加,被试的反应时间变长。而重要的是,在同等的干扰条件下,交叉视差比不交叉视差的反应时间更快,而且目标在干扰方块之前比目标在干扰方块之后的反应时间更快。因此,这个搜索任务的结果也反映了上面所提到的“前方效应”。

2.5 基于双眼视差的立体视觉中的注意过程

在深度维度上是否能形成选择性注意?这曾经是一个有争议的问题。Ghirardelli 和 Folk (1996)认为注意的选择性并不涉及到对深度信息的处理,即注意的“聚光灯”(attentional spotlight)是“深度盲”(depth-blind)的。然而,也已有充分的实验研究结果表明空间注意是能够在深度上分布并具

有“观察者中心”(viewer-centered representation)的倾向,即注意的深度位置离观察者越近,注意的效果(去掩蔽的效果)越高(Andersen, 1990; Andersen & Kramer, 1993; Atchley, Kramer, Andersen, & Theeuwes, 1997; He & Nakayama, 1995; Parks & Corballis, 2006)。Atchley 等人(1997)的知觉负载理论(perceptual load hypothesis)认为,选择性注意出现在深度维度上的前提条件是知觉负载必须充分地高(如有掩蔽干扰的条件下),否则深度维度上的注意就没有必要。

然而, Bauer 等人(2011)研究了在没有掩蔽条件下视觉注意在深度空间维度上的分配。他们让被试辨别立体散点图中有双眼视差方块的深度方向(凸出还是凹入),或该方块的形状(方块还是菱形),并在目标刺激呈现之前的250到600 ms的时间范围内用有效或者无效的线索提示视差方块的深度位置。实验结果表明,无论是深度辨别任务还是形状辨别任务,当目标与线索的间隔时间较短(250 ms)时,在给予有效线索的条件下,被试的平均反应时间要小于无效线索条件下的反应时间。因此,这项实验结果证实了在没有掩蔽刺激的情况下自动性注意也可以在深度空间上分配,而且不受任务类型的影响。

2.6 能量掩蔽和信息掩蔽

在有关感知觉的掩蔽研究中,人们已经逐渐认识到掩蔽作用依据机制的不同,可以分为能量掩蔽和信息掩蔽两个大类(Mattys, Brooks, & Cooke, 2009)。一个掩蔽刺激所造成的掩蔽效果往往含有这两种掩蔽成分。能量掩蔽造成了对目标刺激外周加工的动态反应下降,使进入高级中枢的目标信息有实质性的缺失。信息掩蔽则发生在高级加工阶段,是由于掩蔽刺激和目标刺激有某些相似性时,一些加工资源就会被用于掩蔽刺激,而使得目标刺激和掩蔽刺激之间在高级加工层次上出现竞争与混淆。Wardle 等人(2010)根据实验结果也提出了在他们的双眼去掩蔽范式下所存在的两种掩蔽成分:其中一种成分是不受双眼立体视觉所造成的主观空间分离影响的;而另一部分则可以被主观空间分离所消除。我们认为,第一部分掩蔽属于能量掩蔽的范畴,即不受去掩蔽的知觉线索影响的发生在视网膜水平上的掩蔽,而能被主观空间分离所影响的掩蔽成分属于信息掩蔽,即在知觉加工层次上的掩蔽。

3 双眼视差去掩蔽在科技领域中的典型应用

基于双眼视差立体视觉的去掩蔽效应在日益复杂的科学技术发展以及工程方面获得了越来越多的重视。尤其在图像处理方面,同其它有效线索(如颜色、轮廓、亮度等)一样,双眼视差所形成的深度线索也在图像分割和物体识别操作中也发挥了重要作用(Peterson, Axholt, & Ellis, 2008a, 2008b, 2009; Woo, Kim & Iwadate, 2000)。在视觉搜索任务中,与其它特征线索(如颜色、运动、形状)的利用效果相比较,利用双眼视差的深度线索可以更快和更有效地搜索到目标刺激(Nakayama & Silverman, 1986; Reis, Liu, Havig, & Heft, 2011)。因此,可以将双眼视差的去掩蔽效应应用于立体图像序列的分割算法以及信息交错重叠的复杂环境中以提高目标的可见度(Alper, Hollerer, Kuchera-Morin, & Forbes, 2011; Koch, 1995; Rao et al., 2010)。

视觉显示设备往往传递大量视觉信息,如果处理不当极易造成信号拥挤以致出现信息遮挡、杂乱或者模糊等问题。数据标签(即标记背景中特定物体或者特征的标签,它在空间位置上与其所标记的物体有对应关系)虽然占据较小的屏幕空间,但很容易遮盖物体或者其他物体的标签,并成为信号拥挤的重要因素之一(Peterson et al., 2008a)。在传统的标签算法中,一般通过调整标签的位置来尽量避免标签重叠(Kakos & Kyriakopoulos, 2005; Makita, Kanbara, & Yokoya, 2009; Yamamoto, Gamara, & Lorena, 2002),但由此也带来其它的诸多问题,如难以匹配标签和物体之间的对应关系、标签和物体之间的添加线交叉等等。而更重要的是,非自然标记位置带来的眼动易导致使用者的注意分散(Allendoerfer, Galushka, & Mogford, 2000)。Peterson 等人(2008a, 2009)将双眼视差的去掩蔽作用这一原理用来减少由标签的重叠所带来的视觉信号拥挤,并将这一方法应用到航空监控塔的平视显示器上。在他们的标签分层(label layering)方法中,通过给标签分配不同的视差而将标签分布在若干个深度平面上,从而显著地提高了标签的清晰度和可辨度。研究表明,在标签重叠程度比较严重的情况下,使用有深度线索的标签可以让观察者搜索目标所

用的时间显著地减少(Peterson et al., 2008a)并且不增加不适感或者分心的问题(Peterson et al., 2009)。

鉴于双眼视差显著的去掩蔽作用及其在科技领域的重要应用,我们认为有必要建立和发展双眼去掩蔽的理论计算模型,因为这个模型可以指导图像的获取、编码、显示等过程,即通过分析并调整(增大或缩小)不同深度水平上的物体双眼视差来改变视觉系统所知觉到的物象的立体空间位置,以提高机器图像识别水平并改善人机交互的体验。

此外,如上所述,在复杂刺激场景下,视觉系统对更加靠近观察者深度上的刺激客体给予了更多的加工资源。因此,在用视觉显示设备呈现大量的视觉信息时,可以根据刺激信号的重要性和优先级来调整不同的刺激客体和标签分层,而使得有较高优先级的刺激客体和标签被分配在有较大视差的深度平面上(其视像更靠近观察者),进而使得有较高优先级刺激客体和标签的清晰度和可辨度能有所提高。

4 主观空间分离去掩蔽:大脑处理复杂刺激场景的一个基本功能

综上所述,双眼立体视觉能够形成目标与掩蔽刺激之间在深度维度上的主观空间分离,而这种主观深度分离有助于将注意分配在目标刺激上,进而减少或消除掩蔽刺激的干扰作用。同时,注意在深度维度上的分布有由近到远的权重分布,即更多的注意资源分布到离观察者更近的相对位置上(viewer-centered)。因此,基于双眼视差的立体视觉去掩蔽的效应包括了在心理层次上的信号选择与加工。

有意思的是,主观空间分离去掩蔽现象也出现在听觉加工过程中。当两个相关(correlated)声源(如直达声和其反射声)来自不同的空间位置并且两者间的延迟充分地短,落后声的特征(attributes)就会在知觉层次上被领先声音所“捕捉”(Li, Qi, He, Alain, & Schneider, 2005),而形成在直达声源位置附近的一个融合声像。这种现象被称为优先效应(the precedence effect, 近期研究见 Huang et al., 2011)。在混响环境下,优先效应能产生不相关声音信号间的主观空间分离。例如,当左、右两个扬声器都发出 A、B 两个不同(即不

相关)声音时,对 A 声音来讲,左扬声器提前 3 ms 播放(左右扬声器分别模拟 A 声音的直达和反射声源),而对 B 声音来讲,右扬声器提前 3 ms (右左扬声器分别代表 B 声音的直达和反射声源)。由于优先效应的作用,两个 A 声音和两个 B 声音就分别有知觉融合,其声像分别在左右扬声器处。这样, A、B 声音被知觉为来自不同的空间位置,即出现主观空间分离。但实际上, A、B 声音由两个扬声器所共同呈现,并没有客观上的空间分离。建立在听觉优先效应基础上的主观空间分离并不改变刺激的信噪比,但能够促进听者对目标声音的选择注意进而提高对目标的辨认成绩(Huang et al., 2008; Huang, Huang, Chen, Wu, & Li, 2009; Li, Daneman, Qi, & Schneider, 2004; Wu et al., 2005)。因此,在混响条件下目标信号反射波群与其直达波之间的知觉融合以及掩蔽信号反射波群与其直达波之间的知觉融合,一方面使听者体验不到来自不同方向的反射声,另一方面形成了目标信号和掩蔽信号之间的主观空间分离,进而促进了对目标语音信号的选择性注意。

我们的一个理论假设是,基于双眼视差立体视觉的视觉主观深度分离去掩蔽和基于优先效应的听觉主观空间分离去掩蔽都是建立在以下的一个中枢加工链的基础上:(1)首先,这两种去掩蔽的起始加工都是对刺激的物理精细结构进行编码,而这种编码的精准程度必须充分的高。(2)随后,在对刺激精细结构的编码基础上实现对该精细结构信息之间相关性的中枢计算。对视觉过程来说,这种相关性计算针对两眼之间的信号输入;对听觉过程来说,这种相关性计算发生在领先声音和落后声音(反射声)之间。(3)视觉相关性计算的结果是产生两眼视像之间的知觉融合;听觉相关性计算的结果是产生领先声的声像与落后声的声像之间的知觉融合。(4)对融合后的视像来说,其在深度维度上的位置取决于双眼的视差的性质(正、负)和大小;对融合后的声像来说,其空间位置知觉在领先声源的空间位置附近。(5)如果一个目标视觉刺激与一个掩蔽视觉刺激之间在双眼视差上有所不同,目标视像和掩蔽视像在深度上就有主观空间分离;同样,一个目标声源位置与掩蔽声源位置不相同,目标直达声与其反射波融合的声像的主观空间位置就不同于掩蔽直达声与其反射声的融合声像的主观空间位置(即在目标刺激

和掩蔽刺激之间有主观空间分离)。(6)目标刺激与掩蔽刺激的主观空间分离的结果促进了被试在任务情景下对目标刺激的选择性空间注意,进而提高了对目标的觉察或识别。

此外,视觉的主观空间分离去掩蔽以及听觉的主观空间分离去掩蔽也出现在动物中(Du, J. -Y. Li, Wu, & L. Li, 2009; Du, Wu, & Li, 2010, 2011; van der Willigen, 2011)。

因此,我们认为,形成在知觉加工层次上的目标刺激与掩蔽刺激之间的主观空间分离是大脑在处理复杂刺激场景的一个加工链中的一个重要环节。在这个环节上所形成的主观分离能够促进对目标刺激的选择性注意并压抑与忽略掩蔽信号,进而提高对目标信号的识别。因此,基于双眼视差的立体视觉去掩蔽以及听觉主观空间分离去掩蔽共同反映了大脑在处理大量感觉信息过程中的一个重要的基本功能。

参考文献

- 陈斌, 高闯, 王建中. (2009). 视觉掩蔽研究现状及展望. *心理科学进展*, 17, 1146-1155.
- Adams, D. L., & Zeki, S. (2001). Functional organization of macaque V3 for stereoscopic depth. *Journal of Neurophysiology*, 86, 2195-2203.
- Akerstrom, R. A., & Todd, J. T. (1988). The perception of stereoscopic transparency. *Attention, Perception, & Psychophysics*, 44, 421-432.
- Allendoerfer, K. R., Galushka, J., & Mogford, R. H. (2000). *Display system replacement baseline research report*. Paper presented at Technical Report DOT/FAA/CT-TN00/31, William J. Hughes Technical Center, Atlantic City International Airport (NJ).
- Alper, B., Hollerer, T., Kuchera-Morin, J., & Forbes, A. (2011). Stereoscopic Highlighting: 2D Graph Visualization on Stereo Displays. *IEEE Transactions on Visualization and Computer Graphics*, 17, 2325-2333.
- Andersen, G. J. (1990). Focused attention in three-dimensional space. *Perception & Psychophysics*, 47, 112-120.
- Andersen, G. J., & Kramer, A. F. (1993). Limits of focused attention in three-dimensional space. *Perception & Psychophysics*, 53, 658-667.
- Atchley, P., Kramer, A. F., Andersen, G. J., & Theeuwes, J. (1997). Spatial cuing in a stereoscopic display: Evidence for a "depth-aware" attentional focus. *Psychonomic Bulletin & Review*, 4, 524-529.
- Barlow, H. B., Blakemore, C., & Pettigrew, J. D. (1967). The neural mechanism of binocular depth discrimination. *The Journal of Physiology*, 193, 327-342.
- Bauer, D., Plinge, A., Ehrenstein, W. H., Rinckenauer, G., & Grosjean, M. (2011). Spatial orienting of attention in stereo depth. *Psychological Research*, published online.
- Breitmeyer, B. G. (1980). Unmasking visual masking: A look at the "why" behind the Veil of the "how". *Psychological Review*, 87, 52-69.
- Cartmill, M. (1974). Rethinking primate origins. *Science*, 184, 436-443.
- Changizi, M. A., & Shimojo, S. (2008). "X-ray vision" and the evolution of forward-facing eyes. *Journal of Theoretical Biology*, 254, 756-767.
- Collins, E. T. (1921). Changes in the visual organs correlated with the adoption of arboreal life and with the assumption of the erect posture. *Transactions of the Ophthalmological Society*, 41, 10-90.
- Cottareau, B. R., McKee, S. P., Ales, J. M., & Norcia, A. M. (2011). Disparity-tuned population Responses from human visual cortex. *The Journal of Neuroscience*, 31, 954-965.
- Cottareau, B. R., McKee, S. P., Ales, J. M., & Norcia, A. M. (2012). Disparity-Specific Spatial Interactions: Evidence from EEG Source Imaging. *The Journal of Neuroscience*, 32, 826-840.
- Devisme, C., Drobe, B., Monot, A., & Droulez, J. (2008). Stereoscopic depth perception in peripheral field and global processing of horizontal disparity gradient pattern. *Vision Research*, 48, 753-764.
- Du, Y., Li, J. -Y., Wu, X. -H., & Li, L. (2009). Precedence-effect-induced enhancement of prepulse inhibition in socially reared but not isolation-reared rats. *Cognitive, Affective, and Behavioral Neuroscience*, 9, 44-58.
- Du, Y., Wu, X. -H., & Li, L. (2010). Emotional learning enhances stimulus-specific top-down modulation of sensorimotor gating in socially reared rats but not isolation-reared rats. *Behavioural Brain Research*, 206, 192-201.
- Du, Y., Wu, X. -H., & Li, L. (2011). Differentially organized top-down modulation of prepulse inhibition of startle. *Journal of Neuroscience*, 31, 13644-13653.
- Fox, R., & Patterson, R. (1981). Depth separation and lateral interference. *Perception & Psychophysics*, 30, 513-520.
- Ghirardelli, T. G., & Folk, C. L. (1996). Spatial cuing in a stereoscopic display: Evidence for a "depth-blind" attentional spotlight. *Psychonomic Bulletin & Review*, 3, 81-86.
- He, Z. J., & Nakayama, K. (1995). Visual attention to surfaces in three-dimensional space. *Proceedings of the National Academy of Sciences*, 92, 11155-11159.
- Heesy, C. P. (2009). Seeing in stereo: The ecology and

- evolution of primate binocular vision and stereopsis. *Evolutionary Anthropology*, 18, 21–35.
- Henning, G. B., & Hertz, G. B. (1973). Binocular masking level differences in sinusoidal grating detection. *Vision Research*, 13, 2455–2463.
- Huang, Y., Huang, Q., Chen, X., Qu, T. -S., Wu, X. -H., & Li, L. (2008). Perceptual integration between target speech and target-speech reflection reduces masking for target-speech recognition in younger adults and older adults. *Hearing Research*, 244, 51–65.
- Huang, Y., Huang, Q., Chen, X., Wu, X. -H., & Li, L. (2009). Transient auditory storage of acoustic details is associated with release of speech from informational masking in reverberant conditions. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 1618–1628.
- Huang, Y., Li, J. -Y., Zou, X. -F., Qu, T. -S., Wu, X. -H., Mao, L. -H., et al. (2011). Perceptual fusion tendency of speech sounds. *Journal of Cognitive Neuroscience*, 23, 1003–1014.
- Julesz, B. (1964). Binocular depth perception without familiarity cues: Random-dot stereo images with controlled spatial and temporal properties clarify problems in stereopsis. *Science*, 145, 356–362.
- Kakos, S., & Kyriakopoulos, K. J. (2005). *The navigation functions approach for the label anti-overlapping problem*. In Proceedings of the 4th EUROCONTROL Innovative Research Workshop, Paris, France.
- Koch, R. (1995, 20-23 Jun 1995). *3-D surface reconstruction from stereoscopic image sequences*. Paper presented at the Computer Vision, 1995. Proceedings, Fifth International Conference on.
- Lehmkühle, S., & Fox, R. (1980). Effect of depth separation on metacontrast masking. *Journal of Experimental Psychology: Human Perception and Performance*, 6, 605–621.
- Li, L., Daneman, M., Qi, J. G., & Schneider, B. A. (2004). Does the information content of an irrelevant source differentially affect spoken word recognition in younger and older adults? *Journal of Experimental Psychology: Human Perception and Performance*, 30, 1077–1091.
- Li, L., Qi, J. G., He, Y., Alain, C., & Schneider, B. A. (2005). Attribute capture in the precedence effect for long-duration noise sounds. *Hearing Research*, 202, 235–247.
- Makita, K., Kanbara, M., & Yokoya, N. (2009). *View management of annotations for wearable augmented reality*. Paper presented at the Proceedings of the 2009 IEEE international conference on Multimedia and Expo.
- Manning, M. L., Finlay, D. C., Neill, R. A., & Frost, B. G. (1987). Detection threshold differences to crossed and uncrossed disparities. *Vision Research*, 27, 1683–1686.
- Mattys, S. L., Brooks, J., & Cooke, M. (2009). Recognizing speech under a processing load: Dissociating energetic from informational factors. *Cognitive Psychology*, 59, 203–243.
- McKee, S. P., & Verghese, P. (2002). Stereo transparency and the disparity gradient limit. *Vision Research*, 42, 1963–1977.
- Moraglia, G., & Schneider, B. (1990). Effects of direction and magnitude of horizontal disparities on binocular unmasking. *Perception*, 19, 581–593.
- Moraglia, G., & Schneider, B. (1992). On binocular unmasking of signals in noise: Further tests of the summation hypothesis. *Vision Research*, 32, 375–385.
- Nakayama, K., Shimojo, S., & Silverman, G. H. (1989). Stereoscopic depth: Its relation to image segmentation, grouping, and the recognition of occluded objects. *Perception*, 18, 55–68.
- Nakayama, K., & Silverman, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature*, 320, 264–265.
- O'Toole, A. J., & Walker, C. L. (1997). On the preattentive accessibility of stereoscopic disparity: Evidence from visual search. *Perception & Psychophysics*, 59, 202–218.
- Otto, J. M., Bach, M., & Kommerell, G. (2010). Advantage of binocularity in the presence of external visual noise. *Graefes' Archive for Clinical and Experimental Ophthalmology*, 248, 535–541.
- Parks, N. A., & Corballis, P. M. (2006). Attending to depth: Electrophysiological evidence for a viewer-centered asymmetry. *NeuroReport*, 17, 643–647.
- Patterson, R., Cayko, R., Short, G. L., Flanagan, R., Moe, L., Taylor, E., et al. (1995). Temporal integration differences between crossed and uncrossed stereoscopic mechanisms. *Perception & Psychophysics*, 57, 891–897.
- Peterson, S. D., Axholt, M., & Ellis, S. R. (2008a, 8-12 March 2008). *Managing Visual Clutter: A Generalized Technique for Label Segregation using Stereoscopic Disparity*. Paper presented at the Virtual Reality Conference, 2008. VR '08. IEEE.
- Peterson, S. D., Axholt, M., & Ellis, S. R. (2008b, 15-18 Sept. 2008). *Label segregation by remapping stereoscopic depth in far-field augmented reality*. Paper presented at the Mixed and Augmented Reality, 2008. ISMAR 2008. 7th IEEE/ACM International Symposium on.
- Peterson, S. D., Axholt, M., & Ellis, S. R. (2009). Objective and subjective assessment of stereoscopically separated labels in augmented reality. *Computers & Graphics*, 33, 23–33.

- Preston, T. J., Li, S., Kourtzi, Z., & Welchman, A. E. (2008). Multivoxel pattern selectivity for perceptually relevant binocular disparities in the human brain. *The Journal of Neuroscience*, *28*, 11315–11327.
- Rao, D., Le, Q. V., Phoka, T., Quigley, M., Sudsang, A., & Ng, A. Y. (2010, 18-22 Oct. 2010). *Grasping novel objects with depth segmentation*. Paper presented at the Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on.
- Reis, G., Liu, Y., Havig, P., & Heft, E. (2011). The effects of target location and target distinction on visual search in a depth display. *Journal of Intelligent Manufacturing*, *22*, 29–41.
- Schneider, B., Moraglia, G., & Jepson, A. (1989). Binocular unmasking: An analog to binaural unmasking? *Science*, *243*, 1479–1481.
- Speranza, F., Moraglia, G., & Schneider, B. A. (1995). Age-related changes in binocular vision: Detection of noise-masked targets in young and old observers. *Journal of Gerontology: Psychological Sciences*, *50B*, 114–123.
- Speranza, F., Moraglia, G., & Schneider, B. A. (2001). Binocular detection of masked patterns in young and old observers. *Psychology and Aging*, *16*, 281–292.
- Tsirlin, I., Allison, R. S., & Wilcox, L. M. (2008). Stereoscopic transparency: Constraints on the perception of multiple surfaces. *Journal of Vision*, *8*, 1–10.
- Tsirlin, I., Allison, R. S., & Wilcox, L. M. (2012). Perceptual asymmetry reveals neural substrates underlying stereoscopic transparency. *Vision Research*, *54*, 1–11.
- Tsirlin, I., Wilcox, L. M., & Allison, R. S. (2010). Perceptual artifacts in random-dot stereograms. *Perception*, *39*, 349–355.
- Uttal, W. R., & Fitzgerald, J., & Eskin, T. E. (1975). Parameters of tachistoscopic stereopsis. *Vision Research*, *15*, 705–712.
- van der Willigen, R. F. (2011). Owls see in stereo much like humans do. *Journal of Vision*, *11*, 1–27.
- Wallace, J. M., & Mamassian, P. (2004). The efficiency of depth discrimination for non-transparent and transparent stereoscopic surfaces. *Vision Research*, *44*, 2253–2267.
- Wardle, S. G., Cass, J., Brooks, K. R., & Alais, D. (2010). Breaking camouflage: Binocular disparity reduces contrast masking in natural images. *Journal of Vision*, *10*, 1–12.
- Woo, W., Kim, N., & Iwadata, Y. (2000). *Object segmentation for Z-keying using stereo images*. Paper presented at the Signal Processing Proceedings. WCCC-ICSP 2000. 5th International Conference on.
- Wu, X. H., Wang, C., Chen, J., Qu, H. W., Li, W. R., Wu, Y. H., et al. (2005). The effect of perceived spatial separation on informational masking of Chinese speech. *Hearing Research*, *199*, 1–10.
- Yamamoto, M., Camara, G., & Lorena, L. A. N. (2002). Tabu search heuristics for point-feature cartographic label placement. *GeoInformatica*, *6*, 77–90.

Stereopsis-Based Binocular Unmasking

CHEN Ming-Li¹; ZHANG Chang-Xin¹; YANG Shao-Juan¹; MAO Li-Hua^{1,3};
TIAN Yong-Hong^{2,3}; HUANG Tie-Jun^{2,3}; WU Xi-Hong^{2,3}; GAO Wen^{2,3}; LI Liang^{1,3}

(¹Department of Psychology, Peking University, Beijing 100871, China)

(²School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China)

(³Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing 100871, China)

Abstract: Without affecting the signal-to-noise ratio, the binocular disparity-based stereopsis induces perceived depth separation between the target and the masker (s). The perceived separation improves detection and/or discrimination of the target, leading to the so-called binocular unmasking. This review summarizes previous studies related to binocular unmasking and particularly emphasizes that the allocation of selective visual attention in depth plays a critical role in binocular unmasking. Most importantly, considering that the unmasking effect of perceived spatial separation also occurs in auditory perception, this review proposes that perceived spatial separation-induced unmasking of target signals represents the essential function of the brain in untangling complex scenes. Typical examples of application of binocular unmasking in the industrial field are also provided in this review.

Key words: binocular unmasking; disparity; stereopsis; random-dot stereograms; perceived spatial separation