

La visión en Marr y Berkeley. El problema de perderse el principio de la película

Vision in Marr and Berkeley. The consequence of missing the beginning of the movie

ENRIQUE ARAMENDIA MUNETA*

Resumen: En este artículo se comparan las teorías de Marr y Berkeley sobre la percepción visual a partir de las cualidades establecidas por Descartes (distancia, tamaño, forma y posición). La descripción en tres niveles de Marr, donde la conciencia está ausente, contrasta con el nivel único de Berkeley construido sobre la conciencia. Por otro lado, la experiencia, fundamento último en Berkeley, carece de importancia en los momentos esquemáticos de Marr y cobra protagonismo en el último paso del proceso de la visión mediante la noción de marcación. Bajo la premisa de que las descripciones puramente sincrónicas han de ser forzosamente incompletas, se reformula el viejo problema de Molyneux y se reivindica el papel explicativo de la conciencia en la visión.

Palabras clave: visión, percepción de la distancia, problema de Molyneux, conciencia visual, Marr, Berkeley.

Abstract: In this paper I review Marr's and Berkeley's theories of vision according to four qualities of vision established by Descartes (distance, size, form and situation). I argue that Marr's three level description, where consciousness is absent, clearly differs from Berkeley's one level description, embedded in consciousness. I demonstrate through the notion of markedness that experience gets prominent in the last step of visual processing of Marr's theory of vision, that is, 3D model representation, though it has no role in previous sketchy moments. I also propose a new approach for the ancient Molyneux's problem that deals with diachronic aspects of visual processing and I vindicate the explanatory function of consciousness on vision.

Key words: vision, distance perception, Molyneux's problem, visual consciousness, Marr, Berkeley.

Introducción: ¿qué es lo inmediato?

No es extraño encontrar entre quienes defienden que la percepción visual es algo directo e inmediato ideas divergentes, a veces manifiestamente opuestas, que van desde las formas más sencillas de la psicología de sentido común hasta las más elaboradas y sofisticadas concepciones de intencionalidad¹. Tampoco entonces ha de sorprender que entre quienes niegan esa inmediatez de la visión (por otra parte, el caso más común entre las llamadas teorías clásicas de la visión) se encuentren posturas tan divergentes como las de Descartes, Berkeley

Fecha de recepción: 28/02/2013. Versión corregida: 08/05/2013. Fecha de aceptación: 12/06/2013.

* Universidad Complutense de Madrid. earamendia@ucm.es

1 Para una defensa del realismo directo ajena a la psicología del sentido común, véase Searle (1983, p. 57 y ss).

o Marr. Tal diversidad en torno al concepto de inmediatez requiere empezar con una pequeña revisión acerca de los diferentes sentidos del término, con el doble objetivo de desambiguar su uso en este artículo y asociarlo con legitimidad a aquello que Sellars llamó el *datum*.

Como es bien sabido, Descartes sostenía en la *Dióptrica* (1637, p. 217) que la visión se reduce a seis cualidades básicas, a saber, la luz, el color, la forma (o figura), la distancia, el tamaño (o *magnitude*) y la posición (o *situation*). Todas estas cualidades son perceptibles sin necesidad de la experiencia, puesto que se pueden explicar por medio de una geometría y una física innata. Pero eso no quiere decir que sean perceptibles sin intermediarios por la mente o, dicho en términos cartesianos, por el alma. Para conocerlas, es necesaria la intermediación de las cosas extensas, tales como el cerebro y los filamentos que parten desde el fondo del ojo hacia él².

Para Berkeley únicamente se pueden percibir inmediatamente las dos primeras cualidades: luz y color. Ni la distancia, ni el tamaño ni la posición son percibidos inmediatamente (en el sentido de mediante procesos geométricos o algorítmicos), sino por mediación de otras ideas que sí pueden ser percibidas inmediatamente, es decir, mediante la experiencia. En otras palabras, no existe una predisposición innata para percibir inmediatamente cosas tales como la distancia, el tamaño y la posición, sino un aprendizaje en el que, mediante la asociación de ideas de diferente naturaleza provenientes de distintos sentidos, las cualidades visuales se llegan a percibir.

Marr y Berkeley coinciden en algo: el primer estímulo de la visión es una función de intensidad de luz y color. Ninguna otra cualidad se percibe tan inmediatamente como esas dos. Sin embargo difieren a partir de ese punto. De acuerdo con Marr, la distancia, el tamaño, la posición y la forma se determinan como resultado del procesamiento de esas dos funciones originales.

Todo lo que sucede una vez que la retina ha recogido ese primer estímulo es un proceso cuya descripción requiere de tres niveles: teoría de cálculo, nivel algorítmico/representacional e implementación física³. Durante el proceso de la visión el cerebro computa la función de intensidad de luz recibida en la retina para extraer de ella las partes homogéneas, fundamentalmente bordes, barras y manchas, esto es, los estímulos más efectivos medidos por Hubel y Wiesel (1962) en el córtex estriado. Este es el esquema primitivo. Posteriormente el cerebro computa la orientación y distancia de cada parte en relación al observador junto con sus discontinuidades, lo que dota al esquema primitivo de cierta idea de tridimensionalidad, por lo que se denomina esquema $2\frac{1}{2}D$. Finalmente, el cerebro procesa ese esquema de modo independiente del observador, para reconocer e identificar los objetos gracias a la representación de modelo 3D. Estas tres representaciones que, por motivos que se aclararán más tarde, llamaré *momentos* del proceso visual, pueden derivarse algorítmicamente, es decir, con procedimientos algorítmicos modulares. Sin embargo, Marr niega que el proceso de la visión sea inmediato en el sentido de instantáneo, pues eso equivaldría a minusvalorar

2 *Premièrement, à cause que c'est l'âme qui voit, et non pas l'œil, et qu'elle ne voit immédiatement que par l'entremise du cerveau* (Descartes 1637, p. 224).

3 Para un análisis de la división metateórica de Marr en tres niveles con especial atención al nivel de teoría de cálculo, véase Kitcher (1988).

muchos de los distintos subprocesos que lo componen, cosa por desgracia demasiado frecuente en la bibliografía clásica de la visión.⁴

En suma, cuando Descartes, Berkeley y Marr niegan que la percepción visual sea inmediata se refieren a tres sentidos diferentes de la noción de inmediatez: inmediato como lo que no requiere de intermediarios, inmediato como asociación necesaria o sin mediación de la experiencia e inmediato como lo instantáneo en sentido cronológico. Como uno de los objetivos de este artículo es determinar el papel de la experiencia en la visión y para evitar ambigüedades, en lo sucesivo únicamente se usará el término inmediato en el sentido que empleó Berkeley, es decir, necesario sin mediación de la experiencia pero no por ello necesario sin mediación de otros procesos computacionales o cerebrales.

A continuación se van a comparar dos teorías de la visión tan distantes en tiempo y divergentes en sus cuestiones de fondo como las de Berkeley y Marr. El problema principal en este caso es enfrentar ambas teorías de modo que, como en una prueba de la escala de Mohs, la primera sirva para medir la dureza y consistencia de la segunda, sin caer en el inútil esfuerzo de rehabilitar en pleno siglo XXI un idealismo propio de principios del siglo XVIII. Dada la heterogeneidad de los enfoques, se ha optado por comparar las respuestas de cada una de ellas a temas comunes, para lo cual las cualidades básicas de la visión ofrecen un guión adecuado y riguroso. La distancia o profundidad permitirá esclarecer el papel de la experiencia en el esquema 2½D y dejará clara la necesidad de una descripción diacrónica que complete la explicación puramente sincrónica de Marr. Además, a través de la distancia se comparará el valor explicativo de la conciencia de ambos casos. El tamaño y la forma darán acceso a una nueva revisión del viejo problema de Molyneux, así como a otros aspectos de la experiencia relativos a la noción de marcación, necesaria para la interacción entre derivación y reconocimiento en la representación de modelo 3D. Por último, la posición permitirá perfilar adecuadamente la posición de Marr acerca de la noción de representación y su relación con los procesos computacionales.

La distancia

El principal asunto a la hora de comprender el procesamiento visual de la tridimensionalidad del mundo exterior es la percepción de la distancia entendida como profundidad, esto es, distancia de los objetos con relación al observador. En Marr, la distancia relativa al observador está directamente implicada en la formación del esquema 2½D, esto es, el momento del proceso visual en el que se percibe la distancia subjetiva de lo observado junto con la orientación de su superficie para cada punto de la imagen proyectada en la retina. A efectos explicativos se puede imaginar el esquema 2½D como el resultado de colocar, para cada punto de la imagen retiniana, un pequeño cartón a la distancia y con la orientación calculada en este momento. Por consiguiente en el esquema 2½D no existe todavía un reconocimiento diferenciado de los objetos ni una representación de las caras ocultas.

4 Marr afirma literalmente refiriéndose a los procesos del esquema 2½D lo siguiente: *the term immediate perception is a bit misleading, because these processes can take time.* (Marr, 1982, p. 351)

En el caso de Berkeley, la distancia también se entiende como profundidad (Armstrong 1964, p. 2), aunque se diferencian dos tipos: la idea visible de distancia y la idea tangible de distancia, siendo ambas heterogéneas entre sí. Para evitar la tentación de asociar la idea de distancia de Marr (distancia relativa al observador) con cualquiera de las dos ideas de distancia de Berkeley (idea visible de distancia e idea tangible de distancia), basta tener en cuenta dos aspectos básicos: 1) una distancia relativa al observador puede ser visible y tangible, esto es, puede ser sugerida por la vista o por medición tangible (como sucede al estar «a tres pasos de distancia» de algo) y 2) la distancia relativa al observador, al contrario que las otras dos, es percibida por el sujeto antes de identificar y reconocer los objetos percibidos. Pertenece por consiguiente al momento del esquema $2\frac{1}{2}D$.

Pero aun se puede decir algo más acerca de las nociones de distancia en Berkeley y Marr. Para Descartes la distancia se puede conocer inmediatamente mediante una suerte de geometría natural⁵. Del mismo modo en que un ciego sería capaz de calcular la distancia a la que se cruzan los palos que sujeta con ambas manos conociendo los ángulos que forman y la distancia entre sus manos, la mente es capaz de calcular la distancia con los objetos vistos conociendo la inclinación de los rayos trazados del ojo al objeto y la distancia entre los ojos.

Berkeley, por el contrario, niega que la distancia se pueda percibir inmediatamente. La distancia sólo puede ser percibida por mediación de otras ideas, esto es, por asociación de una idea visible y una idea tangible. Además, ambas ideas de distancia no se pueden relacionar de una manera inmediata, ni tampoco se pueden sumar entre ellas. Sólo se pueden asociar a través de nuestra propia experiencia, gracias a que previamente se ha experimentado de manera simultánea una percepción tangible y otra percepción visible. Por ejemplo, un niño sabe que los elementos más distantes se perciben como más pequeños porque anteriormente, al tratar de coger algo, descubrió que el recorrido hasta tocarlo era mayor cuanto más pequeño se le aparecía.

Marr se encuentra más cerca de Descartes que de Berkeley (Cfr. Kirkeboen, 1998). Como Descartes, Marr cree que es posible derivar geoméricamente, esto es, algorítmicamente, la distancia relativa al observador a partir de la visión. Un poderoso argumento en favor de esta posición reside en el descubrimiento del estereoscopio (Wheatstone, 1838) y los estereogramas de puntos aleatorios (Julesz, 1971), a los que se unen los autoestereogramas (Tyler y Clarke, 1990). No es la intención de este artículo abundar sobre lo ya dicho acerca de estos descubrimientos (véase para ello, por ejemplo, Marr & Poggio, 1979), pero sí recordar lo siguiente:

- a. En general, la estereoscopia demuestra que existe una relación entre la diferencia binocular de un punto observado y la distancia a la que ese punto se percibe. Esto se conoce como estereopsis.
- b. En concreto, el estereoscopio de Wheatstone muestra que es necesario asociar previamente los puntos cuya diferencia binocular se procesará después. Esto se conoce como problema de correspondencia.

5 *Nous connaissons, en second lieu, la distance par le rapport qu'ont les deux yeux l'un à l'autre. Car, comme notre aveugle, tenant les deux bâtons AE, CE, dont je suppose qu'il ignore la longueur, et sachant seulement l'intervalle qui est entre ses deux mains A et C, et la grandeur des angles ACE, CAE, peut de là, comme par une Géométrie naturelle, connaître où est le point E.* Descartes, 1637, p. 222. Ver también fig. 1.

- c. En concreto, los estereogramas de puntos aleatorios de Julesz muestran que la resolución del problema de correspondencia puede darse de manera relativamente independiente a otros procesos visuales y más especialmente de modo previo al reconocimiento de los objetos. Esto se conoce como tesis de la modularidad.
- d. En concreto, los autoestereogramas de Tyler y Clarke sugieren una relación entre la atención y otros factores con la resolución del problema de correspondencia.
- e. En general, los estereogramas muestran que la detección de la distancia puede hacerse de un modo no consciente y por consiguiente que toda descripción de la visión en términos exclusivamente fenoménicos será incompleta.

¿Y cómo se lleva a cabo exactamente esa percepción necesaria de la distancia relativa al observador? La solución de Descartes coincide con la de Marr en que se debe a una especie de *rappor qu'ont les deux yeux*, si bien difiere de ella al no tener en cuenta el problema de la correspondencia así como al usar valores absolutos y no relativos⁶. Para Marr, la percepción de la distancia requiere describir previamente la solución al problema de la correspondencia y después el cálculo de la distancia junto con la orientación de la superficie. Marr describe con detalle la teoría de cálculo y el nivel algorítmico/representacional, pero deja sin respuesta el nivel de implementación física. Además la visión estereoscópica no explica por sí sola cómo se puede tener noción de distancia con un solo ojo. Aun así, la solución de Marr tiene la virtud de enlazar de manera necesaria el cambio fraccional en distancia con la distancia absoluta. A continuación se revisa esta diferencia.

Descartes:

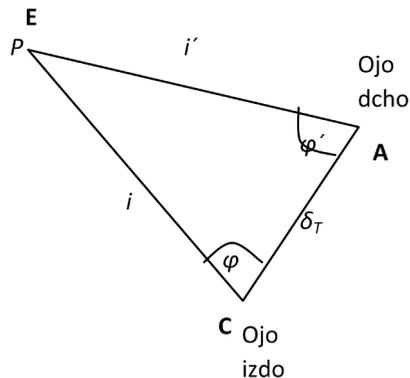


FIGURA 1

Descartes equipara el cálculo de la distancia de un objeto visto al cálculo de la distancia a la que se entrecruzan dos palos que sujeta un ciego con ambas manos. No interesa aquí derivar la fórmula de cálculo, sino subrayar que no contiene distancias visibles, sino una

6 Aquí la palabra «relativo» no indica relación al observador o al objeto, sino *ratio*. Una «relación 1:3» es un valor relativo entre dos distancias, mientras que «dos pies» o «un centímetro» no lo es.

distancia tangible, a saber, la distancia entre ojos, y que esta distancia no es relativa sino absoluta. Junto con los ángulos φ y φ' , la distancia entre los ojos es suficiente para calcular el valor de i , esto es, la distancia a la que se encuentra el objeto con respecto al observador. Por consiguiente el cálculo de Descartes únicamente relaciona distancias tangibles entre sí.

Marr:

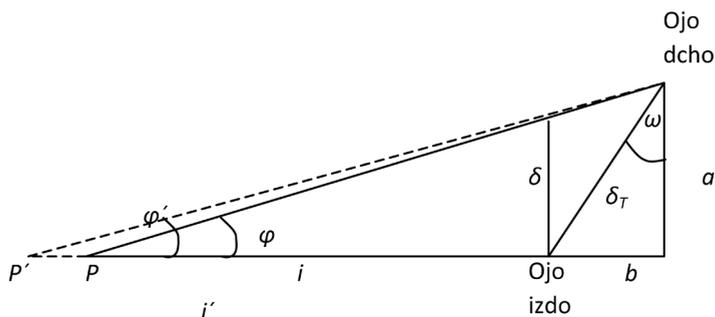


FIGURA 2

$$\left. \begin{aligned} \operatorname{tg} \varphi \approx \varphi &= \frac{a}{i+b} = \frac{\delta}{i} \\ \Delta \varphi = \varphi' - \varphi &= \frac{a}{i'+b} - \frac{a}{i+b} = \frac{a}{i+b} \left(\frac{i+b}{i'+b} - 1 \right) \end{aligned} \right\}$$

Se define q como cambio fraccional en distancia $q = \frac{i'+b}{i+b}$

$$\Rightarrow \Delta \varphi = \frac{\delta}{i} \left(\frac{1}{q} - 1 \right) \Rightarrow \frac{1}{q} - 1 = \frac{i}{\delta} \Delta \varphi$$

Por consiguiente, a través de la disparidad binocular Marr relaciona el cambio fraccional en distancia (q) con la distancia del punto P con relación al observador (i).

Llegados a este punto caben dos interpretaciones diferentes que arrojarán implicaciones distintas para la teoría de la visión en Berkeley. La primera, defendida por Armstrong (1964, p. 5 y ss.), sostiene que la idea inmediata de luz que Berkeley tiene en mente incluye en sí misma la diferenciación entre los ejes x e y . Por consiguiente en ella hay una suerte de orden espacial de dos dimensiones. Si se acepta esa diferenciación, tal y como hacen el propio Marr y Hildreth (1980) cuando postulan la función $I(x,y)$ de intensidad de luz, entonces hemos de asociar el cambio fraccional q a la distancia visible y la distancia i a la distancia tangible. En ese caso, la fórmula de Marr relaciona necesaria e inmediatamente, en el sentido de inmediatez descrito anteriormente, la distancia tangible y visible, en contra de lo que Berkeley postulaba. Este hecho se manifiesta a simple vista en el efecto de deformación de los objetos reconocidos en un estereograma al acercarse o alejarse del mismo,

y que es achacable a que, en ese caso, la distancia visible no varía como correspondería al cambiar la distancia tangible.

La segunda interpretación, defendida por Schwartz (1994, p. 27 y ss.) afirma que la percepción inmediata de la luz a la que hace referencia Berkeley es únicamente una matriz de datos que carece de organización espacial bidimensional. En este caso, la distancia visible y la distancia tangible seguirían sin poder relacionarse de una manera inmediata.

El papel de la experiencia en la percepción de la distancia

La explicación estereoscópica de la distancia no es la única posible ni es la más adecuada en todos los casos. Que no es la única posible lo indica el hecho de que se pueda percibir la perspectiva en un cuadro o de que un tuerto pueda tener una idea aproximada de la distancia de los objetos. Que no siempre es la más adecuada lo indica el hecho de que, a partir de cierta distancia, la diferencia entre lo que ambos ojos perciben es inapreciable (Wheatstone, 1838, p. 380). En el ámbito de la teoría de indicios, Gibson (1950, p. 21 y p. 71-72) contó al menos nueve signos no estereoscópicos que forman parte del estímulo visual y permiten percibir la distancia, entre ellos el eclipsamiento en el campo de visual de los objetos más lejanos por los más cercanos, el cambio de color de los objetos distantes, el acomodamiento del cristalino durante el enfoque de planos a diferentes distancias, el flujo óptico, el cambio de textura y el sombreado. Por su parte, Marr revisa estos tres últimos (1982, p. 212, 233, 239) cuando explica cómo se forma del esquema $2\frac{1}{2}D$, esto es, como se percibe la distancia hasta cada uno de los puntos observados y la orientación local de la superficie tangente a ese punto.

La percepción de la distancia mediante flujo óptico está basada en el llamado paralaje motor (Gibson, 1950, p. 117 y ss.), mediante el cual se relaciona el gradiente de velocidad y dirección del movimiento de la imagen retiniana, apreciable en la deformación del campo visual que percibe un observador en movimiento, con la distancia de las superficies observadas. Gracias al flujo óptico, el observador puede reconstruir una representación fiel de las estructuras tridimensionales que lo rodean y especialmente de su distancia continua, esto es, la orientación con respecto al observador. Pese al enfoque psicofísico original de Gibson, que se basaba en la correspondencia entre estímulos físicos tales como gradientes y experiencias visuales, la percepción de la distancia mediante flujo óptico es perfectamente descriptible en los tres niveles de Marr. Sin embargo, Marr estima que el papel del flujo óptico no es fundamental para los seres humanos, si bien no descarta que pueda serlo para las aves o los insectos voladores. Prueba de ello es el trabajo de Gabbiani, Krapp y Laurent (1999) acerca del proceso de prevención de colisiones en el vuelo de las langostas. Este proceso, que Gibson (1950, p. 127) llamaría detección de distancia por deformación causada por una superficie no paralela a la línea de locomoción, está explicado únicamente mediante descripciones computacionales, algorítmicas y neurológicas.

El cambio de textura es otra forma de percepción de la distancia especialmente adecuada durante la observación de superficies a ras de suelo. El cambio de textura tiene un antecedente en lo que Descartes (1637, p. 224) y Berkeley (1732, ap. 35) llamaron la distinción/confusión y fuerza/debilidad de la figura, en el sentido de que se ocupa de la modificación de esos factores con la distancia así como de cualquier estructura repetitiva a lo largo de

la superficie vista. El cambio de textura (Gibson, 1950, p. 77 y ss.) relaciona con la distancia continua un gradiente, en este caso el gradiente de textura de la imagen retiniana correspondiente a una superficie observada. A cambio, la percepción por cambio de textura requiere que se asuma cierta uniformidad en la estructura superficial y en las condiciones de iluminación de la misma.

Por último el sombreado permite inferir el relieve de la superficie observada a través de las sombras proyectadas por este relieve en la propia superficie. Para ello es necesario suponer un origen o foco de luz sencillo y una función de reflectancia uniforme en la superficie. La ubicación del foco de luz es crítica para la eficacia de este método. Así por ejemplo, dependiendo de lo que se tome por foco de luz, un *cameo* visto con un solo ojo puede parecer *entallado* y viceversa (Wheatstone, 1838, p. 383) y las esferas de un dibujo se pueden percibir como convexas o cóncavas. A este fenómeno se le conoce como inversión del relieve.

Como se puede ver, las explicaciones anteriores requieren de un conocimiento previo (asunción de dirección de movimiento, asunción de uniformidad de estructura, asunción de la ubicación del foco de luz) que no puede provenir sino de la experiencia. Por tanto no debe sorprender que Descartes (1637, p. 224) afirmara que es más fácil percibir la distancia de los objetos que se conocen bien que la de los que se ven por primera vez.

Marr da poca importancia a esos tres mecanismos. Sus descripciones algorítmicas son explicaciones complejas de difícil implementación física y del tipo «foto fija», donde la experiencia no interviene. En otras palabras, la descripción de Marr es sincrónica: explica cómo funcionan los procesos visuales *en el momento*, pero no aclara el modo en que el cerebro ha conseguido desarrollarlos a lo largo del tiempo.

Por el contrario, en Berkeley la experiencia permite asociar los estímulos que se perciben directamente, en su mayoría sensaciones oculares, con la idea tangible de distancia. La naturaleza de esta asociación no es inmediata, sino que se aprende únicamente por experiencia, al modo en que se añaden entradas a una tabla que después será consultada. Este tipo de descripción es diacrónica, en el sentido de que se centra en cómo se adquiere, a través del tiempo, la habilidad descrita. La descripción diacrónica puede admitir también la diferenciación metateórica en tres niveles propuesta por Marr. Dos ejemplos de este tipo de descripción son la hipótesis de sinapsis modificada para las células Purkinje (Marr, 1969, p. 455) o la hipótesis de ajuste de pesos en redes HyperBF a través de las células Golgi (Poggio, 1990, p. 906). Ambas hipótesis describen diacrónicamente la implementación física de respuestas motoras en el córtex cerebelar.

La conciencia en el proceso de la visión

Antes de comparar la noción de conciencia visual en Marr y Berkeley conviene recordar la distinción básica entre conciencia de la descripción de un proceso de visión y conciencia de la visión. En el primer caso, el proceso de la visión es la materia o el asunto que se trae a la conciencia, el cual se puede eliminar sin alterar dicho proceso. En el segundo caso, la conciencia está en el propio proceso y no se puede eliminar sin alterarlo. Cualquiera puede ser consciente de ver, pero no todo el mundo es consciente del modo en que *se ve*. Se llamará a la primera conciencia visual y a la segunda conciencia del proceso visual.

Berkeley (1732, ap. 19) critica la noción de geometría natural e innata que Descartes propone para explicar cómo se percibe la distancia. La percepción de la distancia se lleva a cabo de una manera consciente, sostiene Berkeley. Sin embargo nuestra propia experiencia de percibir la distancia no se parece en nada a un cálculo matemático de líneas y ángulos, ni al ver somos conscientes de estar haciendo una operación geométrica. En consecuencia no es posible que la percepción de la distancia se produzca del modo cartesiano.

La objeción de Berkeley es perfectamente coherente con su idealismo. Toda idea es impresión, proviene de la experiencia y como tal sólo puede ser consciente. Dado que no hay ideas no conscientes, la conciencia visual de distancia y la conciencia del proceso visual de percepción de la distancia deben coincidir, cosa que no se cumple en la dióptrica cartesiana. Ahora bien, ¿es posible aplicar esta objeción al proceso de visión de Marr, dado que éste se apoya también en procesos algorítmicos ajenos a la conciencia? Para responder hay que recordar la descripción de niveles de Marr. Los tres niveles, esto es, teoría de cálculo, algoritmo/representación e implementación física, son descripciones diferentes de un mismo proceso. La existencia de uno de esos niveles no implica la inexistencia de los otros. Al contrario. En el nivel algorítmico Marr trata de explicar aquellos procesos que de manera no consciente se podrían implementar en el nivel físico sobre diversos soportes materiales, de acuerdo con el principio de la múltiple realizabilidad. Para Marr, la geometría innata sólo es la implementación física del proceso de visión. Para Berkeley, por el contrario, la geometría innata es inaceptable porque no hay una implementación física de esa descripción matemática como tampoco hay realidad física fuera de la mente. La conciencia visual es el proceso visual y lo explica al mismo tiempo.

En el caso de Marr, el nivel algorítmico/representacional es únicamente descriptivo y su implementación física no ha de implicar necesariamente conciencia. Tal es el caso, por mencionar sólo un ejemplo, de la detección de cerros en el esquema primitivo (Marr & Hildreth, 1980). El procesamiento neuronal mediante células geniculadas tipo X conectadas con puertas lógicas «AND» es justamente la implementación de un proceso no consciente que previamente se ha descrito en el nivel algorítmico/representacional y en la teoría de cálculo.

Ahora se entiende la importancia de los niveles de teoría de cálculo y algorítmico/representacional para Marr. El análisis en tres niveles de Marr nos permite ser conscientes de cómo funcionan los procesos visuales no conscientes y describirlos adecuadamente al margen de su implementación física.

En suma, la teoría de la visión de Marr, al contrario que la de Berkeley, no rechaza que pueda haber procesos visuales no conscientes, lo cual le confiere cierta ventaja para explicar determinados procesos tales como la detección de bordes o el problema de correspondencia. A cambio, la teoría de la visión de Marr no explica la conciencia visual. Jackendoff (1987) y Prinz (2007) han tratado de rellenar este vacío sosteniendo que es la estructura de nivel intermedio, esto es, el esquema $2\frac{1}{2}D$, la que soporta la conciencia visual. El problema entonces es saber cuál es exactamente la función de la conciencia dentro del proceso de la visión, al menos si no se la quiere convertir en un mero subproducto *ineficiente* de la visión.

El tamaño y el primer momento de la visión

Igual que en el caso de la distancia, Berkeley (1732, ap. 54 y 55) distingue entre idea visible de tamaño e idea tangible de tamaño, siendo esta última a lo que nos referimos cuando comúnmente hablamos del tamaño o magnitud de algo. La idea visible de tamaño es la cantidad de puntos mínimos visibles que se perciben a simple vista, es inconstante (varía de acuerdo a diversos motivos), se percibe inmediatamente puesto que es luz⁷, y no puede ser comparada con la idea tangible de tamaño por tratarse ambas de ideas heterogéneas. La idea tangible de tamaño, por su parte, es la cantidad de puntos mínimos tangibles con las manos, es constante y no puede ser percibida inmediatamente por la vista.

En este caso también se da una discrepancia entre Berkeley y Marr. Berkeley sostiene que la idea visual de tamaño se percibe como *datum*, esto es, inmediatamente, sin necesidad de la intermediación de ninguna otra idea. Asume, por consiguiente, que la mente puede distinguir puntos de diferentes características tales como los que forman contornos que delimitan el tamaño visual. Para Berkeley es un paso natural, dado que esa distinción entre puntos es algo que se produce de manera consciente. Sin embargo James (1890, p. 143 y ss.), sin abandonar todavía el enfoque psicológico, ya advirtió de que la cantidad de terminaciones nerviosas simultáneamente excitadas por un mismo agente exterior no es suficiente para explicar la percepción del tamaño visual, sino que era necesario un cierto orden visual y una cierta capacidad de discriminación.

Marr por el contrario cree que la detección de contornos es uno de los subprocesos más importantes de la visión, cuya complejidad reflejan fenómenos tan cotidianos como el triángulo de Kanizsa o la detección de límites en la piel de un gato o en la superficie estriada del mar. En consecuencia, la percepción de una cantidad de puntos mínimos visibles no implica la capacidad de agruparlos, ni de distinguirlos, ni mucho menos de detectar los bordes que delimitan zonas distinguibles en esos puntos. La negación de esa implicación es precisamente lo que justifica la necesidad de establecer un primer momento del proceso de la visión, esto es, el esquema primitivo, en el cual se derivan los indicadores de lugar (*place token*) que se corresponderán con bordes, barras y manchas (Marr, 1982, p. 51).

Figura, forma y objeto

Otra de las cualidades fundamentales atribuibles a la visión es la figura. La figura implica algo más que la percepción de superficie. Supone la identificación de un objeto percibido al margen del punto de vista del observador, de manera que se puede rotar o mover con la imaginación. Corresponde por consiguiente al momento genuinamente representacional, esto es, a la representación del modelo 3D de la teoría de Marr.

Conviene en primer lugar una advertencia terminológica: para Marr, forma es una cualidad del objeto real, mientras que representación es la imagen formada a partir de lo real. Representación es la descripción (representacional) de una forma (real). De ese modo el concepto más parecido a la figura en Marr es aquel que recoge el conjunto de información

7 Acerca de la idea visual de tamaño percibida inmediatamente se pueden consultar los comentarios de Atherton (1990, p. 115 y ss.)

acerca de la forma real del objeto observado, y que Marr (1982, p. 292) llama genéricamente *elementos primitivos de forma*. Los elementos primitivos de forma son, a su vez, distintos según el momento de procesamiento visual. En los momentos todavía no plenamente representacionales, tales como el esquema primitivo en bruto, el esquema primitivo completo y el esquema 2½D, estos elementos son esbozos en los que aún no se ha llevado a cabo el proceso de identificación de patrones y reconocimiento de objetos. Fundamentalmente son bidimensionales y se denominan elementos primitivos de forma basados en la superficie. Por el contrario, en la representación de modelo 3D los elementos primitivos de forma son tridimensionales, se pueden visualizar desde un sistema de coordenadas relativo al propio objeto y se denominan elementos primitivos volumétricos. Estos elementos primitivos volumétricos son los que permiten reconocer los objetos.

Por desgracia, Marr sólo ofrece elementos primitivos volumétricos basados en conos generalizados, siendo su ejemplo paradigmático la representación de la figura humana. Su insistencia en los conos generalizados es comprensible: son un tipo de elemento primitivo volumétrico inmediatamente sintetizable a partir de una figura del esquema 2½D. Basta con definir un eje y hacer girar la figura sobre él para obtener el cono generalizado, del mismo modo que obtendríamos la forma de un jarrón girando sobre un eje vertical un trozo de alambre moldeado con su perfil. Dicho de otro modo, los conos generalizados tienen *per se* la solución a uno de los más arduos problemas de la representación del modelo 3D: su encaje con el esquema 2½D de una manera necesaria o inmediata, esto es, al margen de la experiencia.

No obstante, Marr no puede evitar otorgar a la experiencia la importancia que requiere en el momento del reconocimiento de objeto. Marr sostiene que dicho proceso se basa en un catálogo de descripciones de modelos 3D, al cual se accede mediante índices de búsqueda que pueden o no ser de origen visual. Este catálogo se alimenta a su vez de la propia experiencia, sea visual o no visual. En otras palabras, si en el esquema primitivo y 2½D la experiencia no jugaba un papel principal, en el caso de la representación del modelo 3D la experiencia se vuelve imprescindible⁸. De ahí la importancia de revisar el modo en que la experiencia determina este último momento y con ello la noción de marcación.

A vueltas con las ideas tangibles y las ideas visibles

Desde el primer apartado de *An essay towards a new theory of vision*, Berkeley distingue entre ideas de la vista e ideas del tacto: distancia visible y distancia tangible, tamaño visible y tamaño tangible, posición visible y posición tangible, extensión visible y extensión tangible. En todos estos casos las ideas tienen una existencia distinta y separada⁹, y sin embargo para designarlas se emplea la misma palabra. ¿Por qué se asocia preferentemente el círculo visible al círculo tangible y no por ejemplo al cuadrado tangible? Para Berkeley, la relación entre las ideas visibles y las ideas tangibles se puede resumir en el problema de Molyneux, que Locke recoge en el capítulo 9 del libro segundo de *An essay concerning human understanding*. Allí

8 *We view recognition as a gradual process that proceeds from the general to the specific and that overlaps with, guides and constrains the derivation of a description of the image.* (Marr, 1982, p. 321)

9 *An existence distinct and separate.* (Berkeley, 1732, ap. 121)

se expone el siguiente caso hipotético: un hombre ciego de nacimiento ha adquirido durante su ceguera cierta destreza para identificar objetos tales como un cubo y una esfera con la sola ayuda del tacto. Si de pronto ese hombre recobrase la vista, ¿sería capaz de identificar la primera vez que los viera aquellos objetos familiares que palpó y tocó cuando era ciego?

El problema de Molyneux es clásico en filosofía. Los abundantes experimentos en pleno siglo XX sobre privación y/o desarrollo de la visión —véase Denegaar (1996, p. 118) para una compilación de experimentos sobre la privación de visión en animales; véase también Hubel & Wiesel (1963), Wiesel & Hubel (1963A y 1963B) y Held & Hein (1963)— reflejan la vigencia de sus cuestiones de fondo, esto es, el papel de la experiencia y la relación intermodal en la percepción, pero también ponen de manifiesto la urgente necesidad de reformular el problema en nuevos términos.

Una de las primeras respuestas al problema de Molyneux fue la de Berkeley, para quien las ideas visibles y las tangibles sólo se pueden asociar mediante la experiencia conjunta de ellas. Por consiguiente para Berkeley no hay una correlación intermodal inmediata entre lo visible y lo tangible, o dicho de otro modo, no hay una relación necesaria entre las variaciones que acontecen en lo percibido por la vista y lo percibido por el tacto.

Durante el siglo XVIII el problema se abordó de modo distinto, especialmente después de que Voltaire (1738, p. 89) hiciese público el caso de un «niño de Molyneux» a quien el cirujano M. Chelselden intervino con éxito de cataratas. Tras la intervención, el niño no fue sometido explícitamente a la prueba de distinguir por la vista un cubo y una esfera, pero se anotó su marcada incapacidad para distinguir formas, tamaños y posiciones, incluso semanas después.

La exposición de Voltaire hizo que por primera vez se viese el problema como un experimento clínico y no sólo mental. En aquella época se diferenciaron tres cuestiones fundamentales:

- a. ¿puede ver el hombre de Molyneux?;
- b. ¿puede distinguir entre ellas la percepción visual de la esfera y la del cubo?;
- c. ¿puede asociar correctamente la percepción visual de cubo a la percepción táctil de cubo y la percepción visual de esfera a la percepción táctil de esfera?

Diderot (1749, p. 132) identificó esas tres cuestiones así como la dependencia mutua entre ellas. Berkeley y Locke habían asumido la cuestión a, esto es, que el ciego sería capaz de reconocer las luces y las figuras, pero que no sería capaz de asociarlas con las ideas tangibles. Sin embargo, Diderot recordó que los ojos del ciego podrían no estar todavía preparados para ver nada más que manchas de luces y colores.

Durante el siglo XIX el enfoque clínico cobró si cabe más importancia, especialmente gracias a los trabajos, entre otros, del doctor J. C. August Franz. Franz es el primero que realizó explícitamente el experimento propuesto por Molyneux, esto es, la distinción entre esfera y cubo, en un paciente operado de cataratas. Franz concluyó que la respuesta a las cuestiones b y c era afirmativa y negativa respectivamente (Denegaar, 1996, p. 93). Sin embargo, hizo otra aportación más. Desglosó el experimento en tres series de pruebas que diferenciaban las habilidades del paciente en la percepción de color, contorno y orientación y finalmente figura. En otras palabras, introdujo nuevas preguntas entre las cuestiones b y c.

El estudio más reciente sobre el problema de Molyneux está basado en pruebas de reconocimiento de formas geométricas en niños ciegos que han recuperado la visión tras

someterse a cirugía dentro del proyecto Prakash (Held, Ostrovsky et al., 2011). Este trabajo avala la respuesta original de Berkeley.

Por lo que a Marr respecta, pese a que no aborde directamente la cuestión, se puede aventurar su respuesta. Para ello es necesario tener en cuenta su división en tres momentos del proceso visual (esquema primitivo, esquema 2½D y representación de modelo 3D). Para los dos primeros momentos, donde aun no tenemos representaciones propiamente dichas sino esquemas dependientes del punto de vista del observador, es probable que Marr diera una respuesta positiva al problema, dado que la experiencia no tiene un papel importante en los procesos allí descritos. Por ejemplo, la detección de ceros que se lleva a cabo sobre la función de luminosidad $I(x,y)$ mencionada antes (Marr & Hildreth, 1980) se puede aplicar a una función de textura en una superficie formada por dos tejidos distintos, de tal modo que obtendríamos igualmente en ambos casos un contorno que indicaría un borde en la forma física. Si hay una descripción matemática idéntica para la percepción de ideas visuales e ideas tangibles, entonces será posible asociar ambas percepciones de una manera necesaria o inmediata.

Esta especulación, en cualquier caso, es cuestionable y aporta poco a la cuestión de cómo se ve (descripción sincrónica del proceso de la visión) y cómo se llega a ver (descripción diacrónica del proceso de la visión). Lo realmente interesante de ella es que pone de manifiesto la necesidad de reformular la cuestión en nuevos términos. El problema de Molyneux exige una reelaboración del experimento que traslade el peso de lo sincrónico a lo diacrónico. Lo importante no es determinar si, justo tras recuperar la vista, el hombre de Molyneux será o no capaz de relacionar las formas percibidas de distintos modos. Lo importante es saber qué sucede hasta que, al cabo de unas semanas y tras un proceso natural de aprendizaje, ese hombre llega a relacionar las formas con éxito.

Schwenkler (2012) ya ha advertido la necesidad de verificar en futuros experimentos la formación completa de la representación 3D antes de concluir la heterogeneidad de las ideas de la vista y el tacto. La propuesta de este artículo es algo distinta. Se trata de reformular completamente el experimento de tal modo que, a lo largo del periodo de aprendizaje, se reproduzcan diversas pruebas cuidadosamente escogidas por su relación con los distintos momentos del procesamiento visual. A continuación se enumeran algunos ejemplos de esas pruebas.

Test	Objetivo
Identificación de la orientación en barras, bordes y manchas (asociar entre sí diferentes percepciones intermodales con igual grado de inclinación)	Poner a prueba la correlación intermodal de la orientación
Identificación de la posición homogénea de puntos en el campo visual	Poner a prueba la posición intermodal relativa al observador
Reconocimiento de movimientos sencillos de barras	Poner a prueba la correlación intermodal de movimiento
Reconocimiento visual de texturas previamente palpadas	Poner a prueba la correlación intermodal de patrones en relieve fino

Reconocimiento de bordes problemáticos (triángulo de Kanizsa, límites borrosos)	Poner a prueba la percepción para el esquema primitivo completo
Reconocimiento visual de patrones en relieve grueso previamente palpados	Poner a prueba la correlación intermodal innata para esquemas 2½D
Reconocimiento de formas tridimensionales por medio del sombreado	Poner a prueba la percepción de formas por medio del sombreado
Reconocimiento de formas tridimensionales por cambio en textura	Poner a prueba la percepción de formas por medio del cambio de textura
Visualización de estereogramas	Poner a prueba la percepción de distancia por estereoscopia (resolución del problema de la correspondencia)

Con esta nueva reformulación se podría por fin saber qué competencias se adquieren antes y cuál es la interdependencia entre ellas, con objeto de describir diacrónicamente los momentos del proceso visual.

La marcación y el reconocimiento

Durante este artículo se han visto las divergencias entre Marr y Berkeley en cuanto a la percepción de la distancia, el tamaño y la figura. Ahora, al tratar el momento de la representación de modelo 3D, se puede hablar de coincidencias. Para ello es necesario detenerse en la noción de marcación.

Como es bien sabido, Berkeley define las ideas abstractas como aquellas ideas que existen al margen de las cualidades particulares asociadas a la experiencia, es decir, con existencia al margen de experiencia. Berkeley niega que haya ideas abstractas. En otras palabras y usando terminología de Locke, no procede la distinción entre cualidades primarias tales como solidez, extensión, forma o movimiento y cualidades secundarias tales como color, olor y sonido, ya que todas las cualidades son secundarias, incluyendo por supuesto la distancia, el tamaño y la posición, y por tanto toda idea proviene de las sensaciones o percepciones particulares.¹⁰ No hay una idea de círculo distinta de las ideas de círculo provenientes de los sentidos. Por consiguiente a través de la idea visible de círculo y la idea tangible de círculo no se puede deducir de manera necesaria una idea abstracta de círculo que englobe a ambas.

¿De qué modo entonces se puede asociar, si es que esto es posible, la idea visible y la idea tangible de un círculo? Para Berkeley la única asociación que cabe entre ambas ideas es la marcación (*markedness*). Esta noción se recoge de manera explícita en el apartado 142 de *An essay towards a new theory of vision*.

I answer, it must be acknowledged the visible square is fitter than the visible circle to represent the tangible square, but then it is not because it is liker, or more of a species with it, but because the visible square contains in it several distinct parts, whereby

¹⁰ Las cualidades secundarias se asocian al modo en que se perciben los objetos, mientras que las ideas particulares son la percepción concreta y particular de esas cualidades en un momento dado.

to mark the several distinct corresponding parts of the tangible square, whereas the visible circle doth not.

Berkeley toma el lenguaje como modelo de marcación. Una palabra escrita no marca correctamente un sonido determinado porque se le parezca, sino porque tiene las suficientes partes para ser asociadas con las partes del sonido. De igual modo, cuando la idea visible de círculo marca la idea tangible de círculo o la idea visible de caballo marca la idea tangible de caballo, no hay un parecido entre ambas ideas. Basta con que las dos tengan el mismo número de partes para que se puedan asociar. Aparte de eso, el único criterio válido para asociar las ideas es la experiencia: gracias a que habitualmente se percibieron al mismo tiempo o de manera asociada, las ideas se marcan unas a otras.¹¹

Esta noción de marcación responde a una lógica de siete premisas:

- a. La marcación de una idea por otra no depende del parecido entre ambas.
- b. La marcación de una idea por otra no se produce por necesidad (4 no marca a 2+2).
- c. La marcación surge tras la experiencia simultánea de dos ideas.
- d. Se requiere de ciertos criterios de adecuación para la marcación.
- e. Las ideas simples son adecuadas para marcar las ideas simples.
- f. Las ideas complejas son adecuadas para marcar ideas complejas.
- g. Una idea compleja es más adecuada para marcar a otra idea compleja si ambas tienen el mismo número de partes.

Las premisas a y b forman lo que se podría llamar el principio de «arbitrariedad de la marcación»¹², básico para comprender cómo funciona el catálogo de ideas de Berkeley. La premisa c, por su parte, enuncia el principio de determinación de la marcación. De la conjunción de las tres premisas se concluye que las ideas tangibles y visibles están asociadas de una manera que, si bien no es necesaria, está determinada en el sentido de que está fijada por la propia experiencia. El catálogo indexado de ideas sensibles se compone de índices cada uno de los cuales se refiere precisamente a esa marcación específica que asocia, en virtud de la experiencia, unas ideas con otras de tal manera que, una vez catalogadas, para identificar dos ideas basta con reconocer una en la otra a través de los índices de búsqueda pertinentes.

Las nociones de marcación y señalamiento son parecidas pero no son equivalentes. Las separa la misma distancia que hay entre la aguja indicadora de velocidad y la marca de la velocidad del cuadro de un coche. La marcación permite recuperar posteriormente lo señalado, mientras que el señalamiento, por sí solo, no permite esa recuperación. Cuando Marr habla de indicadores de lugar en el esquema primitivo no está refiriéndose sino a señalamientos sobre los cuales no es preciso hacer búsquedas puesto que se procesan a partir de la imagen

11 El propio Berkeley (Ibídem) utiliza el paralelismo entre palabras e ideas en los apartados 51, 64, 140, 143, 144 y 147, y es especialmente explícito en este último: *And the manner wherein they signify and mark unto us the objects which are at a distance is the same with that of languages and signs of human appointment, which do not suggest the things signified by any likeness or identity of nature, but only by an habitual connexion that experience has made us to observe between them.* Berkeley también defiende este paralelismo en *Theory of vision vindicated and explained* (1733).

12 No conviene llevar mucho más allá esa noción de arbitrariedad, al menos en el caso de Berkeley: la naturaleza es el lenguaje de Dios y no hay nada arbitrario en ello.

retiniana. En cambio cuando habla de índices en la representación de modelo 3D se refiere a genuinos marcadores que son la base de la interacción entre derivación y reconocimiento, propia del momento de la representación de modelo 3D.

La derivación es la obtención de representaciones 3D adecuadas a partir de los elementos primitivos previos. El reconocimiento, por su parte, es la identificación en esos elementos primitivos de los modelos 3D que han sido derivados anteriormente.

La derivación y el reconocimiento interactúan durante el proceso visual. La derivación de la representación 3D a partir de los elementos primitivos previos del esquema primitivo y esquema 2½D se aprovecha del reconocimiento de descripciones, de tal manera que el propio reconocimiento constriñe la representación que se está derivando. Esta interacción sólo es posible gracias al catálogo de descripciones de modelo 3D. El catálogo está indexado, esto es, *marcado*, con objeto de permitir la búsqueda y relacionar los distintos elementos, tal y como sucedía en el caso de Berkeley. La diferencia, sin embargo, no es trivial. Mientras en el caso de Berkeley los índices provienen de la conciencia y sólo pueden ser simples o compuestos, en el caso de Marr los índices se derivan del proceso visual y su taxonomía depende de esa derivación. Marr define explícitamente tres tipos de índices: los de especificación, los de engarce y los de pertenencia. Esos índices están asociados a elementos primitivos volumétricos, lo que no excluye que haya otros que marquen a partir de elementos primitivos superficiales, tales como patrones de textura o color determinados (las barras blancas y negras, por ejemplo, en el caso de una cebra) o incluso a partir de elementos no visuales (el ladrido, en el caso de un perro).

Del mismo modo, los criterios de adecuación de la marcación en el caso de la representación de modelo 3D no están exclusivamente condicionados por las partes de elementos de igual nivel. Los índices de especificación, por ejemplo, están ligados a la descomposición en elementos más precisos de una representación de trazo grueso. En otras palabras, atienden a una jerarquía de modelos. Por ese motivo las premisas e, f y g de la marcación en Berkeley se sustituyen por criterios de adecuación más sofisticados que tienen que ver con la finura de la representación y la eficiencia del proceso, tales como la estabilidad, la sensibilidad, la accesibilidad y la simplicidad (Marr, 1982, p. 318).

La posición y la dualidad entre proceso y representación

La posición es la cualidad visual por la cual se ubican las partes del objeto observado con respecto al observador y con respecto al resto del campo visual. A la hora de estudiar la posición, Berkeley se interesó especialmente por el problema de la inversión de la imagen retiniana, esto es, por qué se perciben los objetos en su configuración correcta pese a que la imagen proyectada en la retina está invertida. Para Berkeley, el problema de la inversión de la imagen retiniana es para la posición lo que la ilusión de Barrow era para la distancia o el problema de la luna en el horizonte era para el tamaño. Todos ellos son casos donde la explicación geométrica no funciona.

Anteriormente uno de esos geómetras ya había abordado el asunto. En *La Dioptrique* (p. 201), Descartes argumentó que no son los ojos quienes ven, sino el alma, el cual percibe los objetos a través del movimiento que, durante el momento de la visión, se traslada a través de los filamentos del nervio óptico que unen el interior del cerebro con la retina o *fond de l'œil*.

De ese modo, la cuestión acerca de la posición se resuelve como una cuestión mecánica, y el problema visible como un problema tangible. Igual que un ciego es capaz de tocar la parte derecha de un objeto mediante un palo que sujeta con su mano izquierda, nuestra alma es capaz de percibir a la izquierda los rayos de luz proyectados en la parte derecha de la retina.

Berkeley no podía aceptar esta solución. Para él, resolver el problema requiere conocer en profundidad los detalles del mismo. El problema de la posición no es distinguir, por ejemplo, una mano izquierda de una derecha, puesto que la imagen de ambas manos siempre será distinta. Más bien consiste en asociar determinada zona del espacio visible con determinada zona del espacio tangible. Y esto no puede hacerse, como han pretendido los geómetras, de manera necesaria, sino que sólo puede lograrse gracias a la asociación continua de experiencias visibles y tangibles. Así por ejemplo, debido a la configuración celular de la retina, cuando se quiere ver con claridad, esto es, sin confusión, la parte de arriba de un objeto se mueven los ojos en esa dirección. La idea inmediata de claridad, unida a la idea inmediata de movimiento muscular de los ojos, hacen que se perciba la posición de esa parte del objeto. En otras palabras, dos ideas inmediatas permiten percibir una idea mediata como la posición visual, todo ello sin otra mediación que la experiencia¹³. Esa idea visual de posición se asocia a su vez con la idea tangible de posición mediante la experiencia continua y simultánea de tocar la parte superior del objeto y ver la parte superior del objeto.

Marr, al contrario que otros autores como Barlow o Minsky, considera necesario contar con una representación interna de lo visto y defiende que el proceso de la visión consiste precisamente en la construcción de esa representación a través de los distintos momentos del proceso visual, entre los que se encuentran el esquema primitivo y el esquema 2½D. Dado que su teoría es representacional, se podría pensar que Marr debería estar preocupado por el problema de la posición. Sin embargo, no le dedica ni una sola línea. Las razones de esta omisión son más valiosas que la propia omisión en sí: para Marr la representación está relacionada con el proceso. El segundo nivel de explicación del proceso de la visión, que no en vano se llama el nivel de algoritmo y representación, da cuenta de esta relación. Tener una representación adecuada de un objeto observado no equivale a tener una fotografía interna de ese objeto en alguna parte del cerebro. Por el contrario, es el resultado de procesar de la manera eficiente toda la información significativa que la retina recogió hasta obtener una descripción de modelo 3D adecuada para la forma real del objeto observado.

Esa es la dualidad de proceso y representación, en la cual no hay representación sin proceso y ni proceso sin representación. Gracias a que no hay representación sin proceso, la representación no es considerada como un mero calco de la realidad, esto es, como una fotografía en el cerebro, con lo que se eluden problemas tales como el agotamiento de recursos¹⁴ o la falacia del homúnculo. Por otro lado, gracias a que no hay proceso sin repre-

13 Nótese de paso que ese tipo de relaciones entre el movimiento de los ojos y la visión no son negadas por Descartes. Así por ejemplo, Descartes admite que la mente es capaz de percibir el movimiento de los ojos mediante los filamentos del sistema nervioso, de tal manera que a partir de ellos puede conocer los ángulos de visión del objeto (denominados φ y φ' en la figura 1) y finalmente gracias a una suerte de geometría natural, calcular la distancia.

14 Es habitual en la literatura (e.g. Kitcher, 1988; Poggio, 1990) referirse a esa ineficiencia neuronal como la teoría de «neuronas abuela», en la cual cada representación (la de nuestra abuela, por ejemplo) requiere un espacio concreto del cerebro, de tal manera que el mero hecho de tener una representación desde todas las perspectivas posibles agotaría la materia.

sentación se puede mantener una explicación coherente y completa de la visión real en toda su complejidad informacional, sin caer en lo que Marr considera explicaciones fragmentarias de determinados fenómenos de la visión¹⁵, tales como la detección de rasgos de Barlow o la teoría de marcos de Minsky.

Conclusiones

La contraposición de las teorías de la visión en Marr y Berkeley nos remonta hasta el viejo asunto del papel de la experiencia en la percepción y a su expresión en el problema de Molyneux. Ahora se puede reformular el asunto en términos de sincronía y diacronía. La descripción del proceso visual en Berkeley explica cómo se llega a ver a partir de ideas inmediatas y mediante el aprendizaje. Es por consiguiente una descripción diacrónica de la visión. Por otro lado, la descripción de Marr acerca de los momentos esquemáticos (esquema primitivo y 2½D) y del momento representacional 3D es una descripción sincrónica, que explica cómo se ve *en el momento*, fotograma a fotograma, una vez que la habilidad visual ha sido correctamente adquirida. En esta explicación nos falta el principio de la película, es decir, cómo se ha llegado a desarrollar esa habilidad. Esta perspectiva exclusivamente sincrónica no es infrecuente, se encuentra en muchas explicaciones funcionalistas y también en la respuesta de Held y otros (2011) al problema de Molyneux. En este artículo se propone una reformulación diacrónica del problema través de pruebas concretas que permitan dirigir la atención del experimento al proceso de adquisición de la capacidad visual.

Pese a todo, no es justo decir que Marr renuncia completamente a la experiencia. La noción de marcación, imprescindible para entender la interacción entre derivación y reconocimiento de la representación de modelo 3D, es la prueba fehaciente de la importancia de la experiencia en el proceso visual. Dado el carácter forzosamente incipiente de esa parte del trabajo de Marr, era de suponer que con el tiempo la experiencia y el aprendizaje habrían de adquirir mayor importancia en la descripción del proceso visual. Esto ya se puede constatar en los trabajos posteriores de Poggio (2004).

El otro gran tema que surge de contraponer estas dos teorías de la visión es la conciencia visual. En Berkeley, la conciencia visual y la conciencia del proceso visual coinciden, la conciencia explica la visión y es *datum*, esto es, no se explica mediante otras ideas. En Marr, la conciencia no explica el proceso visual, sino que es un asunto más, cuya explicación se puede retrasar indefinidamente. Entre la explicación circular de Berkeley (conciencia como puro *explanans*) y la paciente moratoria de Marr (conciencia como puro *explanandum*) es necesario dejar hueco para una nueva fenomenología de la conciencia.

15 Dos muestras de esto: 1) lo que dice Marr (1982) acerca de Minsky y otras aproximaciones desde la inteligencia artificial en p. 347: *The problem is that the goal of such studies is mimicry rather than true understanding, and those studies can easily degenerate into the writing of programs that do not more than mimic in an unenlightening way some small aspects of the human performance.* 2) lo que dice Marr (Ibídem) acerca de la aproximación de Barlow a la detección de rasgos en p. 341: *The visual description necessary to solve this problem have to be more complex and less direct related to what we naturally think of as their representation as a string of motor strokes.*

Referencias

- ARMSTRONG, D. M. (1960). *Berkeley's theory of vision: A critical examination of bishop Berkeley's essay towards a new theory of vision*. Melbourne University Press.
- ATHERTON, M. (1990). *Berkeley's revolution in vision*. Cornell University Press.
- BERKELEY, G. (1732) *An essay towards a new theory of vision (4th ed.)*. En LUCE A. A., JESSOP T. E. (Eds.), *The works of George Berkeley, bishop of Cloyne. Vol. one: Philosophical commentaries. An essay towards a new theory of vision. Theory of vision vindicated*. London: Thomas Nelson and Sons Ltd.
- DEGENAAR, M. (1996). *Molyneux's problem: Three centuries of discussion on the perception of forms* (M. J. Collins Trans.). Dordrecht, Netherlands y Boston: Kluwer Academic Publishers.
- DESCARTES, R. (1637). *La Dioptrique*. En *Oeuvres et lettres*. Editions Gallimard. (1953) [*Discurso del Método, Dióptrica, Meteoros y Geometría*. Ediciones Alfaguara. Madrid. 1981.]
- DIDEROT, D. (1749). *Lettre sur les aveugles a l'usage de ceux qui voient*. En P. Vernière (Ed.), *Oeuvres philosophiques de Diderot* (1964). Paris: Garnier Frères.
- GABBIANI, F., KRAPP, H. G., & LAURENT, G. (1999). Computation of object approach by a wide-field, motion-sensitive neuron. *The Journal of Neuroscience*, 19(3), 1122-1141.
- GIBSON, J. J. (1950). *The perception of the visual world*. Cambridge, Massachusetts: The Riverside Press.
- HELD, R., & HEIN, A. (1963). Movement-produced stimulation in the development of visually guided behavior. *Journal of Comparative and Physiological Psychology*, 56(5), 872-876.
- HELD, R., OSTROVSKY, Y., GANDHI, T., GANESH, S., MATHUR, U., & SINHA, P. (2011). Corrigendum: The newly sighted fail to match seen with felt. *Nature Neuroscience*, 14(9), 1217-1217.
- HUBEL, D. H., & WIESEL, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, 160(1), 106-154.
- HUBEL, D. H., & WIESEL, T. N. (1963). Receptive fields of cells in striate cortex of very young, visually inexperienced kittens. *Journal of Neurophysiology*, 26(994), 1002.
- JACKENDOFF, R. (1987). *Consciousness and the computational mind*. (1994) The MIT Press.
- JAMES, W. (1890). *The principles of psychology. Volume two (1950)*. New York: Dover Publications.
- JULESZ, B. (1971). *Foundations of cyclopean perception*. The University of Chicago Press.
- KIRKEBØEN, G. (1998). Descartes' psychology of vision and cognitive science: The optics (1637) in the light of Marr's (1982) vision. *Philosophical Psychology*, 11(2), 161-182.
- KITCHER, P. (1988). Marr's computational theory of vision. *Philosophy of Science*, 55, 1-24.
- MARR, D. (1969). A theory of cerebellar cortex. *The Journal of Physiology*, 202(2), 437-470.

- MARR, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. Cambridge, Massachusetts. The MIT Press. 2010. [*La visión*. Alianza Editorial. Madrid. 1985.]
- MARR, D. & HILDRETH, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 207(1167), 187-217.
- MARR, D. & POGGIO, T. (1979). A computational theory of human stereo vision. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 204(1156), 301-328.
- POGGIO, T. (1990). A theory of how the brain might work. *Cold Spring Harbour Symposia on Quantitative Biology. Vol. LV*. 899-910.
- POGGIO, T., & BIZZI, E. (2004). Generalization in vision and motor control. *Nature*, 431(7010), 768-774.
- PRINZ, J. (2007). The intermediate level theory of consciousness. En M. Velmans, & S. Schneider (Eds.), *The Blackwell companion to consciousness* (pp. 248-260) Blackwell Publishing.
- SCHWARTZ, R. (1994). *Vision: Variations on some berkeleyian themes*. Blackwell Publishing.
- SCHWENKLER, J. (2012). On the matching of seen and felt shape by newly sighted subjects. *I-Perception*, 3(3), 186-189.
- SEARLE, J. R. (1983). *Intentionality: An essay in the philosophy of mind* Cambridge University Press.
- SELLARS, W. (1956). *Empiricism and the philosophy of mind, intro by Richard Rorty and study guide by Robert Brandom* (1997) Cambridge: Harvard University Press.
- TYLER, C. W., & CLARKE, M. B. (1990). Autostereogram. *SPIE Proceedings, 1256* (Stereoscopic Displays and Applications) 182-197.
- VOLTAIRE. (1738). *Éléments de philosophie de Newton, divisés en trois parties*. En *Oeuvres complètes: Physique. Philosophie de Newton. Tome vingt-troisième*. (1818). París: Plancher.
- WHEATSTONE, C. (1838). Contributions to the physiology of vision. Part the first. On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society of London*, 128, 371-394.
- WIESEL, T. N., & HUBEL, D. H. (1963A). Effects of visual deprivation on morphology and physiology of cells in the cat's lateral geniculate body. *Journal of Neurophysiology*, 26(978), 6.
- WIESEL, T. N., & HUBEL, D. H. (1963B). Single-cell responses in striate cortex of kittens deprived of vision in one eye. *Journal of Neurophysiology*, 26(1003), 17.