

# VISIÓN POR COMPUTADORA

## Tarea 1

### Reporte de lectura: Haciendo que las máquinas (y la inteligencia artificial) vean

Ángel García Báez  
*zs24019400@estudiantes.uv.mx*

Maestría en Inteligencia Artificial

**IIIA** Instituto de Investigaciones en Inteligencia Artificial  
Universidad Veracruzana  
*Campus Sur, Calle Paseo Lote II, Sección 2a, No 112*  
*Nuevo Xalapa, Xalapa, Ver., México 91097*

20 de febrero de 2025

## 1. Introducción

Este ensayo da un breve resumen, así como comentarios específicos acerca de la lectura Haciendo que las maquinas (y la inteligencia artificial) vean Hurlbert (1988). Al final también se dará una breve conclusión acerca de la lectura.

El presente ensayo busca dar un resumen de los puntos clave acerca de la lectura Haciendo que las maquinas (y la inteligencia artificial) vean Hurlbert (1988) y haciendo observaciones puntuales sobre la perspectiva del autor en dichos apartados.

## **2. Cuerpo**

### **2.1. La visión como objeto de estudio**

La visión como objeto de estudio (Antes de poder ver, que hay ahí) Estamos acostumbrados fuertemente a ver sin cuestionarnos como es que lo hacemos, podríamos atribuir que la vista es una mera percepción y que el cerebro se encarga de todo lo demás, pero puede que no sea del todo así. En una historia de intentos y discusiones a lo largo de las décadas, se a intentado definir la inteligencia por si misma sin llegar a un punto solido para poder definirla. De la mano con esto, el articulo propone que la visión es más que un sentido mediante el cual entra información a nuestro cerebro, si el cerebro (o lo que ocurre dentro del cerebro) se le puede considerar inteligente ¿Por qué no incluimos a la visión en esta definición? Aquí es donde se liga el concepto de ver con la IA, puesto que una de las características que debería tener un ser inteligente, es la capacidad de ver para actuar en su entorno (se incorpora aquí la robótica para generar el movimiento). Sonaba aparentemente fácil “solo haz que la computadora vea y listo” ella sabrá interpretar las cosas, más temprano que tarde, los investigadores se percataron que lo que aparentaba ser un problema trivial que podía resolverse en un verano, termino por convertirse en toda una línea de investigación que vigente hasta el presente año 2025 (y seguirá por muchas décadas más). Bajo la hipótesis tradicional del sistema físico de símbolos, era cuestión de que la computadora lograra “ver” extraer información de lo que veía como símbolos para hacer manipulaciones que eventualmente la llevaran a poder describir lo que veía, el problema saltaba a la vista: ¿Cómo hacer eso? ¿Cómo calzar la visión con algún tipo de sistema experto que hiciera las inferencias lógicas sobre lo que “veía”? Aquí fue cuando se dieron cuenta que al nivel que lo querían manejar, la visión no es algo trivial para el ser humano y mucho menos para una máquina.

### **2.2. Visión de bajo y alto nivel**

Si retomamos la idea de que un ser humano es en si mismo, un manipulador de símbolos actúa y llega a conclusiones lógicas para producir comportamiento inteligente, entonces, ¿Por qué la visión no se basa enteramente en un esquema similar? La respuesta puede radicar en el hecho de que esa información que percibimos mediante los ojos, es convertida y representada de una manera distinta a la que entra para ser procesada por el cerebro, entonces, si la cosa apunta por ahí, es necesario convertir la entrada a una que la computadora pueda entender. Nosotros vemos quizás una

imagen y nos es de lo más natural, pero la computadora lo que esta viendo, es una matriz (arreglo bidimensional) de pixeles, en donde cada casilla solo hay números. ¿Cómo hacer para que la maquina entienda que un conjunto de números ordenados es un objeto? Basados en la forma en como los ojos humanos procesan lo que ven, se llega a la identificación del reconocimiento de bordes como una de las tareas primarias (visión de bajo nivel) que hacemos antes de que si quiera sepamos que es lo que tenemos enfrente. Como siguiente paso, se encuentra la “visión de alto nivel”, en la que después de la identificación de bordes cuenta con procesos más avanzados como lo es el reconocimiento de color, textura, profundidad, etc en las representaciones a partir de imágenes “planas”. Finalmente, un hilo conductor que une ambos pasos, es la existencia de “restricciones naturales”, es decir, la identificación de situaciones que no pueden ser plausibles, como que una hormiga sea más grande que un elefante, eso seria un caso ilógico de llegar a conjeturar por la maquina, dado el conocimiento del mundo con el que cuenta, por ello, la implementación de restricciones que delimiten los casos posibles que puede conjeturar la maquina en su ejercicio de “ver” el mundo apuntan a la mejora de su “interpretación”.

### **2.3. Dos perspectivas: IA clásica y el conexionismo.**

Una vez visitado todo lo que se encuentra de fondo, visto parte de la historia del desarrollo para llegar hasta este momento en donde se le busca dotar de la cualidad de ver a la computadora es que salta la pregunta ¿Cuál es el camino correcto?.

Por un lado, el enfoque de la IA clásica sostiene que esto se puede abarcar desde la manipulación simbólica mediante la creación de un sistema experto con cientos de reglas enfocadas a la visión de objetos. El problema de esto sucede al percatarse de la existencia de leves perturbaciones en las imágenes, como las que se producen en nuestros propios ojos si tratamos de observar al mismo objeto con cada uno por separado, dicho sistema de reglas se ve sobrepasado en cuanto el objeto que esta tratando de catalogar en base a su poderosa inferencia, no cumple con los criterios para ser identificado. Por ejemplo, asumiendo que exista un conjunto de reglas que mapea en una imagen a un teléfono de casa de color negro, se sabe que dicho teléfono tiene un color específico, una forma y una altura a la que se encuentra. Si se cambia el color del teléfono de negro a blanco, en automático el sistema hiper especializado en detectar al teléfono se ve incapaz de detectar al teléfono blanco que aparece en escena pese a que tenga la ventaja de ser explicable, al saber que es lo que hace la maquina por detrás.

Por otro lado, el enfoque de la IA conexionista, el cual se caracteriza por creer que la inteligencia emerge a partir de unidades simples de redes interconectadas. Propone que es necesario emular la forma en como descomponemos nosotros, en nuestros ojos y cerebro a las imágenes mediante arreglos bidimensionales que puedan ser introducidos y procesados por una computadora. Dicho procesamiento no es arbitrario, pues la idea de este enfoque es poder enseñarle a la computadora a que aprenda por si misma el cómo descomponer una imagen en bordes para que detecte objetos y posteriormente, pueda darle una interpretación pero con la particularidad de que esto se hace mediante la interconexión de redes que se encarguen de toda la tarea, llegando así a algún resultado pero sin saber cómo llegaron ahí ni que hay en medio del proceso.

Esto ultimo, pese a que parece que encamina mejor el tema de la visión, tiene más de un problema: Es computacionalmente costoso el procesamiento de una sola imagen, a esto se le suma el hecho de que para emular la visión, es necesario procesar más de una imagen por segundo junto con el hecho de que se esta creando una red que detecta patrones pero no es explicable.

En medio de estos dos enfoques que persiguen lo mismo, es donde se sitúa el científico de la visión por computadora.

### 3. Conclusión

La visión por computadora es un área extensa que no se limita unicamente al procesamiento de grandes matrices de datos (imagenes) como pensaba, si no que es un área que va más allá de lo computacional para apoyarse en los modelos fisico-quimicos de como es que los humanos vivos vemos, para poder encapsular todo este conocimiento para que una maquina tenga la capacidad de percibir el mundo mediante imagenes, extraer información sobre el y actuar en consecuencia.

### Referencias

Hurlbert, Anya y Poggio, T. (1988). Making machines (and artificial intelligence) see. *Daedalus*, pages 213–239.