

Latencias de la imagen estadística

Roc Albalat

BAU, Centro Universitario de Artes y Diseño de Barcelona (España)

<https://dx.doi.org/10.5209/TEKN.95149>

Recibido: 19 de marzo de 2024 • Aceptado: 09 de julio de 2024 • Avance en línea: 17 de octubre de 2024, OPR

URL del vídeo: <https://vimeo.com/924287718?share=copy>

ESP Resumen. El ensayo explora las genealogías que relacionan los actuales sistemas de visión por computadora y de generación de imágenes mediante inteligencia artificial con los usos de la fotografía en las investigaciones científicas del siglo diecinueve. Tomando como casos de estudio las imágenes obtenidas por Étienne-Jules Marey y Francis Galton, el audiovisual fomenta el análisis de las latencias o mutaciones culturales que manifiestan las imágenes operativas de ambas épocas.

Palabras clave: aprendizaje automático; fotografía; IA generativa; iconografía; inteligencia artificial.

ENG Latencies of the statistical image

EN Abstract. The essay explores the genealogies that link today's computer vision and artificial intelligence imaging systems to the uses of photography in nineteenth century scientific research. Taking the images of Étienne-Jules Marey and Francis Galton as case studies, the audiovisual encourages an analysis of the latencies or cultural mutations that the operative images of both periods manifest.

Keywords: Artificial intelligence; generative AI; iconography; machine learning; photography.

Sumario. 1. Texto. 2. Declaración de uso de LLM. 3. Referencias.

Como citar: Albalat, Roc (2024) Latencias de la imagen estadística, *Teknokultura. Revista de Cultura Digital y Movimientos Sociales*, avance en línea, 1-2. <https://dx.doi.org/10.5209/tekn.95149>

1. Texto



Los actuales sistemas de visión por computadora y de generación de imágenes mediante inteligencia artificial (IA) tienen genealogías históricas particulares. Ambos sistemas se basan en automatizar procesos de extracción de datos a través de imágenes. Unos procesos que, a una escala menor, ya se habían puesto en práctica durante el siglo diecinueve, cuando la cultura

tecnocientífica encontró en la fotografía un medio con el cual registrar y medir los patrones invisibles del mundo. El vídeoensayo ‘Latencias de la imagen estadística’ tiene como objetivo explorar estas genealogías mediante dos casos de estudio: por un lado, la técnica cronofotográfica del fotógrafo e investigador Étienne-Jules Marey y, por el otro, los retratos compuestos del estadístico Francis Galton. El formato audiovisual del ensayo busca propiciar un análisis iconológico, haciendo manifiestas las formas que emergen tanto en las imágenes fotográficas victorianas como en las actuales imágenes de la IA y facilitando la comprensión de las latencias o mutaciones culturales de la imagen operativa.

La imagen fotográfica del siglo diecinueve se ve alentada por una cultura científica interesada en capturar los gestos humanos más comunes. Las cronofotografías suponían para Étienne-Jules Marey un método de inscripción y síntesis gráfica para el análisis del movimiento, en el cual la imagen facilitaba la conversión del movimiento en datos. La reducción de los detalles de la imagen a una huella de líneas esquemáticas facilitaba el análisis de la posición de los cuerpos y objetos durante un lapso temporal determinado (Marey, 1894). Con más de un siglo de diferencia, la mirada que opera en este tipo de imagen se reproduce en los actuales automatismos de visión artificial. Las redes neuronales profundas de estimación de posturas, como Alpha Pose, nos devuelven imágenes concomitantes con las que obtuvo Marey: la síntesis gráfica del cuerpo que se imprime en las cronofotografías coincide con las formas de la salida gráfica de la herramienta de detección. Sin embargo, hay una diferencia notable en la función que cumple la imagen. La simplificación de las formas le procura a Marey un diagrama, un dispositivo a escala perceptiva humana, para poder analizar lo que en directo resulta demasiado rápido y confuso. Por el contrario, en las imágenes del algoritmo Alpha Pose la solución gráfica ya no es una condición operativa. Los datos de posición de los cuerpos, mediante la indexación de puntos clave, quedan inscritos en un archivo de datos JSON, mientras que la visualización esquemática de la detección de los cuerpos es una imagen posterior a la captura de estos datos y cumple una función deferente con el ojo humano. La imagen sintética no es necesaria para las funciones de monitoreo de los cuerpos, las cuales pueden realizarse de forma automática y en directo, sin la intervención de la mirada humana.

Los retratos compuestos de Francis Galton, o *pictorial statistics* según su propia nomenclatura (Galton, 1883), muestran las medianas fisionómicas de una colección de rostros sobreimpresos sobre la misma placa fotográfica. Los rasgos alineados con la mediana se perciben oscuros, mientras que las desviaciones conforman zonas borrosas en la imagen. Mediante estudios como este, los estadísticos positivistas buscaban establecer patrones objetivos de ‘normalidad’, a partir de los cuales aplicar ratios de desviación. Para Galton, el aspecto de la imagen resultante conformaba una «composición ideal» –hoy la entendemos como la representación de unos ideales elitistas y eugenistas–. Cabe remarcar que, en los rostros compuestos de Galton, la cámara fotográfica no está al servicio del realismo documental, sino que se emplea más bien como un dispositivo de fusión. Mediante una compilación de recortes, consigue reunir un conjunto de imágenes de caras para luego combinarlas, en lo que resulta una imagen de imágenes. La red neuronal StyleGAN2 constituyó, a principios de la presente década, un hito en la generación de imágenes de aspecto fotorrealista. Ante un archivo digital de decenas de miles de retratos, la máquina busca soluciones estadísticas óptimas de píxeles que también puedan formar parte de dicho archivo. Los rostros que genera el modelo son un reflejo de las características más comunes del conjunto de datos (*dataset*) con el que se ha entrenado, reproduciendo ciertas características prevalentes, tal como ocurría con las zonas oscuras de los compuestos de Galton. Cuando observamos la exploración de los rostros consecutivos que genera la red neuronal, vemos representadas las medianas probabilísticas del *dataset* original. La principal diferencia con las imágenes de Galton radica en la evidencia del sesgo. Mientras que los retratos compuestos estaban indexados mediante taxonomías, haciendo evidente, qué incluye y qué omite la mediana estadística, en las imágenes de StyleGAN2, confeccionadas mediante un archivo mucho mayor –y generalmente oculto por imperativos industriales–, resulta muy difícil conocer cuál es la mirada que construye el modelo. Las características específicas del archivo de imágenes con el que la red se ha entrenado quedan codificadas y puestas a funcionar en adelante.

2. Declaración de uso de LLM

Este vídeo no ha utilizado ningún texto generado por un LLM (ChatGPT u otro) para su redacción.

3. Referencias

Galton, Francis (1883). *Inquiries into human faculty and its development*. MacMillan Co.
Marey, Étienne-Jules (1894). *Le mouvement*. G. Masson.