CC BY

ASRI
ARTE Y SOCIEDAD
REVISTA DE INVESTIGACIÓN EN ARTES Y HUMANIDADES DIGITALES

CONTRA EL MONTAJE: SÍNTESIS Y FLUIDEZ VISUAL EN EL VIDEOCLIP PRODUCIDO CON INTELIGENCIA ARTIFICIAL

AGAINST EDITING: SYNTHESIS AND VISUAL FLUIDITY IN THE MUSIC VIDEO PRODUCED WITH ARTIFICIAL INTELLIGENCE

Ana Sedeño-Valdellos
Universidad de Málaga
Rafael Marfil-Carmona
Universidad de Granada

DOI: 10.33732/ASRI.992

Recibido: (10 05 2026)
Aceptado: (22 07 2026)

Cómo citar este artículo
Sedeño-Valdellos, Ana y Marfil-Carmona, Rafael. (2026). Contra el montaje: síntesis y fluidez visual en el videoclip producido con inteligencia artificial. *ASRI. Arte y Sociedad. Revista de investigación en Arte y Humanidades Digitales,* (31), e992.
Recuperado a partir de https://doi.org/10.33732/ASRI.992

Resumen
La modificación del universo de la imagen que supone la llegada de herramientas de generación automática ha alcanzado todas las fases y ámbitos de su producción, distribución y consumo, afec-

Abstract
The transformation of the visual landscape brought about by the advent of automated generation tools has permeated all phases and areas of production, distribution, and consumption, impacting the

tando transversalmente al mundo cultural. El videoclip, como formato de experimentación visual, recoge estas nuevas posibilidades para aplicarlas a las necesidades de la industria musical y del modelo de ocio del siglo XXI. Este trabajo tiene como objetivo describir las características generales de estos procesos de finalización del video musical producido con IA, señalando sus líneas emergentes en lo relacionado con los códigos de montaje/edición y postproducción. Se ha estudiado una muestra de videoclips, aplicando criterios de análisis de contenido, centrados en la dimensión textual y discursiva, con especial interés a los efectos de la producción con IA sobre su montaje. Los resultados apuntan a que los clips de la muestra contienen diversas tácticas visuales que enmascaran el montaje entendido como cambio de plano. Se trata de una tendencia a la transición por forma o composición —*match-cut*—, la unión de *frames* como versiones o iteraciones por modificación de *prompts*, así como a efectos como el *morphing*. La consecuencia es una edición donde destaca la fluidez visual y una estética de síntesis.

Palabras clave
Inteligencia artificial, videoclip musical, postproducción, estética audiovisual, montaje.

cultural world across the board. Music videos, as a format for visual experimentation, embrace these new possibilities and apply them to the needs of the music industry and the 21st-century leisure model. This work aims to describe the general characteristics of these AI-generated music video completion processes, highlighting emerging trends related to editing and post-production techniques. A sample of music videos was studied, applying content analysis criteria focused on the textual and discursive dimensions, with particular attention to the effects of AI production on editing. The results indicate that the clips in the sample contain various visual tactics that mask editing understood as a change of shot. These include a tendency towards transitions through form or composition —match-cuts— the joining of frames as versions or iterations through modified prompts, and effects such as morphing. The result is an edition that stands out for its visual fluidity and a synthetic aesthetic.

Keywords
Artificial intelligence, music video, postproduction, audiovisual aesthetic, editing.

Introducción: inteligencia artificial, industria musical y videoclip

La imagen contemporánea es una imagen factual, procesal y que siempre desea estar en comunicación con el espectador. Su orientación es ser aceptada y consumida, despertar el interés de los receptores. Es persuasiva y performativa por naturaleza. Sus procesos se desarrollan en paralelo a las trayectorias tecnológicas de sus medios de producción, que unen a una lógica tecnológica y a la exigencia de máxima innovación. Por ello, las imágenes proliferan y circulan, en la cultura contemporánea, junto y gracias a formatos audiovisuales que responden y experimentan con nuevos procedimientos creativos. Siguiendo con la dinámica del protagonismo del *software*, como bien apunta Manovich, la inteligencia artificial se afianza como el sistema técnico predominante en la producción visual actual. Las redes neuronales adversarias, los *transfers* de estilo y modelos de lenguaje, las redes convolucionales y los algoritmos de *machine learning* dominan hoy la generación de imágenes (Gómez-Diago, 2024; Morgan, 2021; Ortiz Arroba, 2024), y realizan una redefinición del proceso creativo para una mayor automatización y eficiencia (Jiménez-Jacinto et al., 2025). Con ello permiten múltiples usos como la restauración y mejora, el reconocimiento de partes y tareas y, sobre todo, la automatización de su producción. Este último aspecto es clave en la actualidad, y tiene consecuencias en una

mayor eficacia en la eliminación de artefactos o ruido, en el almacenamiento, así como en la identificación de elementos (cuerpos, rostros...) (Bruccoleri et al., 2022). Esto conduce a la optimización de la imagen con fines económicos e institucionales y tiene consecuencias sobre las tareas de los profesionales y los flujos de trabajo conocidos tradicionalmente (Cazzaniga et al., 2024; Franganillo, 2022).

Las imágenes creadas por estas técnicas de inteligencia artificial generativa profundizan en su performatividad (Sarzi-Ribeiro, 2023) por lo que no se quedan en su propia producción, sino que generan nuevos vínculos entre productores, espectadores y agentes del mercado cultural. El cambio en la dimensión creativa influye en todo el proceso comunicacional. Desde que la llamada IA generativa (GenAI o IAgen) ha llegado a la producción visual se reconoce un salto cualitativo. No son escasos los estudios empíricos y reflexiones sobre sus rasgos y tendencias y el debate sobre su estética, tratada desde 2018 por Lev Manovich. En su más reciente *Estética artificial: IA generativa, arte y medios visuales* (Arielli y Manovich, 2025), se explora cómo la IA generativa está transformando la creatividad, convirtiéndose en agente productor, pero siguiendo una senda basada en estilos visuales anteriores, cuyas características son filtradas mediante algoritmos para reconocer patrones y aplicarlos en nuevos contenidos (Jiang et al., 2022).

A pesar de que la IA aplica modelos anteriores —como dispositivo estético basado en última tecnología—, aparenta responder a un proceso novedoso. Manovich (2018; Arielli y Manovich, 2025) denuncia esto y demuestra que en tanto la IA es entrenada mediante todo el acervo visual histórico, su lógica estratégica se mantiene en torno a las operaciones básicas del bricolage, la cita y la intertextualidad surgidas en las vanguardias artísticas de principios de siglo XX. El resultado, en su aire de nuevo modelo de creación, es propio de la denominada sociedad *posdigital*.

Los desafíos que ocupan estas nuevas imágenes se multiplican. Crawford se pregunta, con un hálito incómodo, qué política hay detrás de la IA, a qué intereses responde y cómo debería limitarse atendiendo a sus consecuencias sobre la totalidad de lo visual (Crawford, 2022, p. 44) en lo relacionado con su tendencia a la homogeneidad figurativa. Efectivamente, existe cierta limitación en los estilos que producen y priman lógicas de collage de las etapas DIY *(Do it yourself)* de internet (Mazzi, 2023). Para Schober (2024), también componen un conjunto de mayor uniformidad que las imágenes creadas por los humanos, tanto que los retratos generados por IA se ajustan a tropos visuales muy normativos, con mirada frontal, composición simétrica e iluminación idealizada. Atendiendo a esta perspectiva general, se comprueba cómo la IA transforma todos los aspectos creativos contemporáneos, alcanzando también los relacionados con la industria musical, especialmente sensible en su necesidad de innovación continua, con las consecuencias de una mayor creatividad, autonomía, propiedad intelectual, inversión, desarrollo tecnológico, flexibilidad y escritura musical (Lis Gutiérrez y Pulido-Flórez, 2024, p. 95). Es la tensión creativa la que impulsa el mercado.

Teniendo en cuenta este contexto de transformación, este trabajo se centra en abordar el videoclip musical creado con IAgen para realizar una primera aproximación en lo relacionado con sus códigos de montaje, es decir, profundizando en el aspecto puramente textual y discursivo, sin perder de vista todo el cambio que va generando el mundo de la industria audiovisual y musical. Se trata, por tanto, de una aportación profesional para la comprensión de códigos, desde la perspectiva del análisis para la comprensión crítica y la clarificación de procesos creativos. Se prevé, en cualquier caso, la tendencia a la experimentación con estas herramientas, lo que será claramente observable a corto plazo. De hecho, trabajos como los de Caballero (2023), López (2024) y Chen et al. (2022) describen estas innovaciones.

Universidad Rey Juan Carlos
ASRI - Arte y Sociedad. Revista de Investigación
Núm. 31. Año 2026. ISSN: 2174-7563
CUVPAC Grupo de investigación CULTURA VISUAL & PRÁCTICAS ARTÍSTICAS CONTEMPORÁNEAS

Los sistemas de inteligencia artificial como Kaiber, Capcut, Adobe Premiere Pro (con Sensei) o Final Cut realizan sincronizaciones automáticas, detección de planos y escenas (Lee, 2022; Andrews, 2023; Smith, 2022) y recomendación de mejores momentos para la edición y sincronización automáticas (Kumar, 2023). Esto produce una delegación de decisiones en el algoritmo, que acoge nuevos roles de agencia, pero también aumenta la automatización de procesos, que no siempre son transparentes para el humano.

Objetivos, materiales y método

En el trabajo se propone un análisis exploratorio de videoclips que permita reflexionar sobre las modificaciones que la aplicación de inteligencia artificial está teniendo sobre el montaje/edición de este género, asumiendo que deben afectar a su estética y estar relacionados con la historia del formato, combinando lo viejo y lo nuevo.

Como rasgo básico de la realización audiovisual, el montaje construye una representación temporal que va más allá de la ordenación de unidades y la linealidad. Así, el videoclip, como máximo exponente de la estética de postproducción de Nicolas Bourriaud (2009), ha demostrado una sensibilidad especial para desarrollar sus posibilidades, a la vez que usa técnicas especialmente complejas, fruto de la era de la composición digital, que se han convertido en estándares de los efectos incluidos en postproducción de cualquier formato audiovisual.

Esta construcción textual sigue lógicas de experimentación y creatividad que retoman todo el acervo de la tradición e historia del audiovisual incluso en su etapa *postelevisiva* (Sedeño-Valdellós, 2025; Sedeño-Valdellós et al., 2016), lo que se traduce en una mayor libertad discursiva adecuada para el desarrollo promocional de la música popular. En este trabajo se tiene en cuenta la reflexión sobre el montaje que ha asistido a la teoría del cine, en un nuevo escenario digital donde sigue siendo válida la definición de Jacques Aumont, considerando el montaje como "la herramienta mental" que administra "la relación entre lo visible e invisible" (2015, p. 33).

En términos metodológicos, se ha realizado una búsqueda bibliográfica y documental sobre la inteligencia artificial en formatos audiovisuales, las tendencias estéticas y la creatividad para contextualizar el trabajo, lo que se ha abordado durante el marco teórico. De igual manera, se ha indagado sobre los códigos de postproducción y sus —probables— cambios por la inteligencia artificial generativa: en general, se puede resumir en términos de finalización, mejora y automatización de procesos, cuando no de una nueva dimensión que concierne también a su estética (Anantrasirichai y Bull, 2022; Lee, 2022; Jironza Hidalgo, 2024; Caballero, 2023; Chen et al., 2022). En esa línea de trabajo, se trata de una aportación que combina la revisión documental con el estudio muestral, profundizando en una cuestión absolutamente técnico-lingüística y discursiva.

El universo de videoclips producidos con inteligencia artificial es imposible de controlar en su totalidad debido al volumen de producción y los diversos canales donde se comparten. La muestra se abordó según la tradición del formato, que necesita de una gran comprensión e indagación previa, con una búsqueda general de piezas o videoclips realizados con inteligencia artificial, lo que se conoció a través de artículos de divulgación, blogs o páginas como las que se citan a continuación. Esto terminó con la localización de 39 casos, cuyos datos han sido obtenidos desde sitios web especializados, como *Top of the clips* (2023),

o de listas de reproducción en canales temáticos de YouTube, Vevo y otras plataformas de la industria musical. Antes de abordar la discusión en torno a los resultados y las conclusiones, se han incluido los resultados en la Tabla 1, a modo de resumen y compilación, incluyendo los elementos más relevantes.

Tabla 1: Muestra de videoclips analizados.

Videoclip / Artista / Año	*Link*
Agua pa ti. Locomía. 2024	https://www.youtube.com/watch?v=aSzKVMFc1S8
The Streechers. *Paintress.* 2021	https://www.youtube.com/watch?v=LnznY1s-tLI
Die Antwood. *Age of illusion.* 2022	https://www.youtube.com/watch?v=Cq56o0YH3mE
Disturbed. *Bad man.* 2022	https://www.youtube.com/watch?v=fpUpVznI4Yc
Hora Zulú. *Alegorías del gran Misterio.* 2023	https://www.youtube.com/watch?v=A_MZgZhkacU&t=63s
Periphery. *Atropos.* 2023	https://www.youtube.com/watch?v=Ppg8kpG-lio
Linkin Park. *Lost.* 2023	https://www.youtube.com/watch?v=7NK_JOkuSVY
La Juventud. *Tremendo.* 2021	https://vimeo.com/420611304
Bazarian. *Can We pretend.* 2021	https://www.youtube.com/watch?v=bSrorge6QWM
Kasabian. *Scripture.* 2022	https://www.youtube.com/watch?v=R2BmQZPrTgk
FKA Twigs. *Pamplemousse.* 2022.	https://www.youtube.com/watch?v=A9gcHsFkiek&list=RDA9gcHsFkiek&start_radio=1
Washed Out. *The Hardest Part.* 2024	https://www.youtube.com/watch?v=-Nb-M1GAOX8
TSHA. *Running.* 2024	https://www.youtube.com/watch?v=TfjoxzwU96k
Josman. *AhGars!* 2023	https://vimeo.com/788588457
Duran Duran. *Danse Macabre.* 2023	https://www.youtube.com/watch?v=BpWrSTBP5rg
London Grammar. *Higher.* 2023	https://www.youtube.com/watch?v=mOKopwtRdb0
Jacques. *Absolve.* 2024	https://vimeo.com/874986396
Cordelia Gartside. *16xo.* 2023	https://www.youtube.com/watch?v=PcIZ7SGD7gg
Ab-Soul & Kendrick Lamar. *Herbert.* 2022	https://vimeo.com/816714697
Cores. *Bacará.* 2023	https://vimeo.com/826902471
Skinny Pelembe. *Deadman Deadman Deadman.* 2023	https://www.youtube.com/watch?v=_AEWTBBhlBg
Jonathan Wilson. *Marzipan.* 2023	https://www.youtube.com/watch?v=jxzeOELVjkc
Xanguito. *Tot.* 2023	https://www.youtube.com/watch?v=vfwOoeFCmio
Billy Joel. *Turn the lights back on.* 2024	https://www.youtube.com/watch?v=UOf6CMbHPuA
Noah Bouchard. *Sometimes.* 2024	https://www.youtube.com/watch?v=4qYsA0jGbQg
Iceboy Vilet & Nueen. *Closer.* 2024	https://www.youtube.com/watch?v=q26RE7TCZBY
Biznaga. *La gran renuncia.* 2024	https://www.youtube.com/watch?v=7t6jdfH3dDU&t=42s
Paul White. *Surreal.* 2024	https://www.youtube.com/watch?v=OOiTllkxN1o&t=58s
Visions of Hell. *Unjust Memes.* 2024	https://www.youtube.com/watch?v=VmYBq9O5774&t=1s
Fatboy Slim & Daniel Steinberg. *Bus Stop Please.* 2025	https://www.youtube.com/watch?v=UDFr7JRe7os
Kanye West. *Bomb.* 2025	https://www.youtube.com/watch?v=qeWutl38TQw
RL Grime & 070 Shake. *Pour Your Heart Out.* 2023	https://vimeo.com/838700437
Ivan Llopis. *Blurred Cliff.* 2025	https://vimeo.com/1000695075
PNAU, Troye Sivan. *You Know What I Need.* 2022	https://www.youtube.com/watch?v=G05cC2xasOY
Björk & Rosalía. *Oral.* 2023	https://www.youtube.com/watch?v=8jsi2Tgvx6A
Desert. *Que vol aquesta gent.* 2023	https://www.youtube.com/watch?v=AwsRdDgUfd4
¥$, Ye, Ty Dolla $ign, Bump J & Lil Durk. Vultures (Havoc Version). 2024	https://www.youtube.com/watch?v=GQEcxrY0CWA
Grimes. *Artificial Angels. 2025*	https://www.youtube.com/watch?v=tvGnYM14-1A&t=41s
Queen. The Night Comes Down. 2024	https://www.youtube.com/watch?v=jiNVqhtjyW4

Fuente: Propia.

Universidad Rey Juan Carlos
ASRI - Arte y Sociedad. Revista de Investigación
Núm. 31. Año 2026. ISSN: 2174-7563
CUVPAC Grupo de investigación CULTURA VISUAL & PRÁCTICAS ARTÍSTICAS CONTEMPORÁNEAS

Resultados: montaje del videoclip musical producido con inteligencia artificial

El primer desafío para caracterizar los videoclips que se están produciendo con IA tiene que ver con la generación de imágenes por iteración a través de *prompts* —*text to image* o *text to video*—: el análisis demuestra que la generación mediante *prompt* de imágenes o vídeos es una técnica predominante, presente en 35 videoclips, que tiene muchas consecuencias en la estética y el montaje del formato. La primera y más importante se concreta en la generación de un flujo visual por suma de imágenes, en que termina por convertirse el montaje: interesante es apuntar cómo este procedimiento recuerda a la técnica de interpolación de fotogramas o *keyframes*, técnica clave en animación. Este efecto, que sobresale tras el visionado de casi todas las piezas de la muestra, subraya una relación de estas con la era de la postproducción y etapas anteriores del audiovisual digital (Korsgaard, 2017; Kumar et al., 2020).

En este sentido, puede hablarse de una similitud visual con técnicas básicas de animación como el *stop motion*. Diversos modelos de IA contienen formas de predecir y generar material entre dos fases o situaciones: así podría citarse a LCM-Lora, con la que —una vez generada una primera imagen con un primer *prompt*— la herramienta proporciona una interfaz para bocetar la composición y realizar múltiples versiones o iteraciones que unidas entre sí produce este flujo visual en continuidad. Por otro lado, Generative Session de Runaway, por ejemplo, permite cargar una imagen y conseguir otras muchas similares. En un paso más allá, los últimos modelos de IA como Krea.AI permiten dibujar formas que son traducidas sirviéndose de *prompts* e integradas en tiempo real en el flujo de trabajo de vídeo. En cualquier caso, el modelo IA devuelve material similar que forma esa cadena de *frames* continuos, en cuyo desarrollo se encuentra, posiblemente, una mayor libertad creativa que en la acción humana del montaje tradicional, en tanto se produce de manera aleatoria.

Cuando se reflexiona sobre sus consecuencias en la estética clip se localizan algunas anomalías reseñables, como supone el fenómeno de que la mayoría de los ejemplos de tipología performance se alejen del tradicional espacio de concierto en exterior o interior donde aparece la banda. La escenografía como elemento narrativo, por tanto, no se circunscribe a la actuación, sino que busca espacios más simbólicos. Por el contrario, la realización aprovecha las variaciones de formas y líneas para construir una destacada riqueza visual, lo que desdibuja bordes de lugares y formas, y se aparta de la representación figurativa del *performer* con fondo convencional: esto ocurre en *Pamplemouse* de FKA Twigs o *You know what I need* de PNAU y Troye Sivan, pero también en otros casos como *Deadman Deadman Deadman* de Skiny Pelembe, que aporta casi 500 rostros generados por DALL-E (Top of the clips, 2023), sobre todo tipo de fondos que se mencionan.

Este mismo efecto tiene aplicación especial sobre la dimensión física de las figuras humanas, ya sea sobre su rostro o su cuerpo. Esto se produce gracias a la unión de imágenes estáticas generadas y de su suma en el flujo de montaje para formar lo que se denomina en teoría visual unión por forma o *match-cut*. Este tipo de montaje, herencia de las primeras vanguardias históricas experimentales, fue teorizado por Eisenstein (1991) como un tipo sintáctico de asociación de imágenes, recurso de continuidad consistente en ensamblar dos escenas o planos a través de similar forma/composición.

En esa línea, se aprecian varios ejemplos que apuestan por la mezcla o juego entre ilustración o dibujo/imagen fotográfica, proponiendo una fusión perceptiva y, de nuevo, mayor fluidez, ayudada por posiciones

Universidad Rey Juan Carlos
ASRI - Arte y Sociedad. Revista de Investigación
Núm. 31. Año 2026. ISSN: 2174-7563
CUVPAC Grupo de investigación CULTURA VISUAL & PRÁCTICAS ARTÍSTICAS CONTEMPORÁNEAS

y gestos similares de las imágenes: *Pamplemousse* de FKA Twigs o *Higher* de London Grammar (Figura 1) son ejemplos de ello. Estas imágenes constituyen un palimpsesto visual donde cada *output/frame* generado es parte del flujo.

Figura 1: Arriba: London Grammar. *Higher*. 2023. Capturas de videoclip.

Abajo: FKA Twigs. *Pamplemouse*. 2022. Capturas de videoclip.



Fuente: Arriba: https://www.youtube.com/watch?v=mOKopwtRdb0.

Abajo: https://www.youtube.com/watch?v=A9gcHsFkiek&list=RDA9gcHsFkiek&start_radio=1.

Una transmutación del cuerpo *performer* aparece en muchas piezas, y ello se relaciona con cómo la identidad construida visualmente atraviesa el objetivo promocional del videoclip. En *Lost*, de Linkin Park, se comprueba una transformación corporal fluida y una fusión con elementos naturales como el fuego, el agua, estrellas o galaxias, y una conversión del rostro del protagonista en armadura robótica. Estas partes de la performance se caracterizan por una continua transformación visual, a través de variadas formas de transición generativa unidas a efectos fractales y de pixelación, *glitch*, y ruido visual.

En este sentido, tanto la imagen como los cuerpos contenidos en ella parecen desmaterializarse produciendo en la banda imagen efectos de disolución que aumentan esa percepción de fluido visual y camuflan el cambio clásico de plano. Esto acentúa la dirección poética de la letra de la canción, que trata de momentos de debilidad, confusión o falta de dirección vital, como evidencia el mismo título *Lost* (Perdido) y desarrollan algunas frases "*cause when I am alone, I lost in these memories, living behind my own illusion, lost of my dignity, living behind my own confussion*" ("Porque cuando estoy solo, Estoy perdido en estos recuerdos. Viviendo detrás de mi propia ilusión. Perdí toda mi dignidad. Viviendo dentro de mi propia confusión").

En varios momentos, el cuerpo se confunde y fusiona con objetos de diversa materialidad, se pierde, se desmaterializa en la nube de puntos 3D, en las burbujas del mar, en el ruido espectral televisivo y se convierte en un flujo de transmutación sin fin, donde parece que la única continuidad es el propio cambio.

Otro ejemplo en esta línea es el videoclip de *Atropos* de Periphery: una fábula en animación tradicional 2D con una fotografía que recrea un mundo oscuro y solitario con una mezcla de paisaje urbano apocalíptico, donde solo carreteras, coches y edificios enormes son visibles en la ciudad, junto a la representación de seres solitarios que caminan con la mirada fija en su celular. Los movimientos automáticos y antinaturales de estos seres parecen querer visualizar un cuerpo y una identidad en deconstrucción por la modificación de rostros, de nuevo conseguidos a través de la unión de *prompts* iterativos.

En relación con ello, otros procedimientos son visibles, especialmente en el componente performance: parece emplearse la inteligencia artificial para animar imágenes estáticas y construir transiciones imperceptibles entre ellas, simulando encadenados casi perfectos o directamente *morphings* (modificaciones visuales graduales que aplicadas a los rostros parecen transfiguraciones): esto ocurre en *Aqua pa ti* de Locomía, de manera muy evidente, pero también en *Tot*, de Xanguito. En este sentido, hay que destacar que, si la esencia del montaje era la imposibilidad de animar la imagen artificialmente, apostando por la combinación para comunicar un sentido narrativo, el *software* de IA resuelve la posibilidad de contar toda una historia desde una sola imagen fija que adquiere movimiento.

Otros modos de camuflar los cambios de plano aparecen en variadas piezas de la muestra, que, en general, se podrían observar como variantes de estas anteriores: en *Que vol aquesta gent* de Desert, el paso de unas formas visuales a otras se produce en una mayor gradación, de manera que las diversas escenas o piezas se trasmutan unas en otras por su cercanía en encuadre (Figura 2): la cabeza de la chica en ojo, el ojo en tronco de árbol... Son saltos entre zonas del encuadre, no combinación de planos diferentes, con la posibilidad de creación de planos detalle que no se producen por registro sino por generación sintética/aleatoria.

Figura 2: Desert. *Que vol aquesta gent*. 2023. Capturas de videoclip.



Fuente: https://www.youtube.com/watch?v=AwsRdDgUfd4.

Las posibilidades de postproducción se convierten también en recurso en este afán de creación de flujos visuales en las imágenes, de manera que el trabajo de composición digital o por capas se emplea tanto

Universidad Rey Juan Carlos
ASRI - Arte y Sociedad. Revista de Investigación
Núm. 31. Año 2026. ISSN: 2174-7563
CUVPAC Grupo de investigación CULTURA VISUAL & PRÁCTICAS ARTÍSTICAS CONTEMPORÁNEAS

para integrar unos tipos de imágenes en otras como para lograr un movimiento en profundidad, como si se estuviera viajando en el espacio de la imagen. A ello ayuda la capacidad de la IA para copiar estilos visuales —transfer de estilo—, lo que además permite asegurar continuidad o *raccord* visual. La facilidad de los agentes artificiales para simular figuraciones y estilos de todo tipo es acogida dentro de la estética del videoclip, forjada ya en operaciones de copia, en el pastiche y la intertextualidad. Esta se manifiesta con una tendencia hacia una *figuratividad* de ilustración y dibujo, ya sea con animación tradicional (2D) o tridimensional (3D). Videoclips como *Bad Man* de Disturbed, *Danse Macabre* de Duran Duran, *Periphery* de Atropos, *Age of Illusion* de Die Antwood o *Que vol aquesta gent* de Desert (Figura 3) explicitan cómo el dibujo animado se convierte en figuración 2D, aunque destaca también cómo se incluye este como capa en entornos híbridos, donde convive en encuadre con otros tipos de imágenes en *Ah Gars* de Josman, *Herbert* de Ab-Soul & Kendrick Lamar, o en el intermedio del flujo de imágenes (*You Know What I Need* de PNAU y Troye Sivan, *Deadman Deadman Deadman* de Skinny Pelembe, *Higher* de London Grammar o *Pamplemouse* de FKA Twigs).

Figura 3: Atropos. *Periphery*. 2023. Captura de videoclip.



Fuente: https://www.youtube.com/watch?v=Ppg8kpG-lio.

El cómic animado y directamente el anime no son mayoritarios en la muestra, pero sí manifiestamente una opción que sobresale y se observa en un mismo primer visionado: el estilo Ghibli requiere mención particular, lo que resulta coherente con diversos fenómenos de viralidad en internet, como han demostrado algunos estudios (Vásconez Merino et al.,, 2025). Estas opciones parecen estar asociadas a una estética emocional y nostálgica y se relacionan con el protagonismo cultural del videojuego y la estética japonesa (Deckker y Sumanasekara, 2025).

Tiene similares resultados que otra aplicación visual muy recurrente, el efecto *morphing*, y que puede visionarse claramente en *Tot* y *Vultures*, donde algunas transiciones dejan -de nuevo- un resquicio de

pixelización y *datamoshing*: el forzar la unión de estos *frames* similares construye una forma de montaje automático, que, paralelamente, camufla el montaje por el aumento de la fluidez, junto a una especie de vibración o parpadeo visual. Este resulta característico también, según se ha comprobado (Zhao y Ye, 2021; Vear y Poltronieri, 2022), de la estética de otros formatos producidos con IA.

Otros efectos visuales derivados, que conducen fácilmente a esta estética del devenir visual, llegan desde la capacidad de síntesis y transformación digital de la imagen. Así ocurre con las herramientas que simulan la visualidad de las aplicaciones para escaneado 3D y estratigrafía, convirtiendo el material en nube de puntos al que se añade una especie de cámara virtual que consigue una flotabilidad automatizada sobre el espacio de la imagen. Esta especie de movimiento de *travelling* surge, en realidad, mediante simulación, gracias a la capacidad generativa: es decir, no es un registro convencional sino un efecto de la forma en que la tecnología "sintetiza" o crea las imágenes. En este territorio visual el espacio se vuelve orgánico, navegable por la cámara, que se desliza en profundidad y en todas direcciones: el espectador percibe moverse por el espacio representado, al igual que ocurre en el mundo visual de un videojuego. Y esto produce una nueva concepción casi de la puesta en serie.

De igual manera, en búsqueda de esta continuidad visual, el director Paul Trillo hizo 55 clips con Sora y los editó en Adobe Premiere (Noesfm, 2024) para Washed Out en el clip *The Hardest part*, toda una declaración de la puesta en valor de la fluidez visual a través de *zoom outs*, con resultado de explosión desacelerada de los elementos icónicos. Se reconoce aquí la senda del *one shot music video* o video de plano secuencia, un patrón observado en gran producción en los últimos diez o quince años, una especie de modo-video en su etapa *postelevisiva* (Sedeño-Valdellos, 2025) donde primaba también la continuidad visual frente al montaje por cambio de plano clásico. También en *Pour Your Heart Out*, las texturas musicales electrónicas se visualizan a través de *zoom ins* cuya duración son las frases musicales en los estribillos; los *zooms outs* o movimientos de cámara dentro del espacio de la acción o traspasan la materialidad de cuerpos u objetos gracias a técnicas como el *depth mapping* o mapeo de la imagen, que hace navegable su espacio. Ello da lugar a momentos donde la *figuratividad* de la imagen se pone en cuestión, más interesada en visualizar el grano del material o directamente generar efectos similares al *glitch* o *noise*. Todos estos últimos ejemplos remiten de nuevo a las capacidades de generación de la IA que priman la fluidez y la continuidad. Se comprueba un efecto de homogeneización de recursos y simulación de figuraciones o estilos de imagen.

Conclusiones

Este trabajo aporta una descripción e interpretación exploratoria de una muestra que, en su esencia, representa una clara tendencia de los videoclips musicales generados con IA. En este sentido, se han identificado nuevos recursos técnicos y expresivos vinculados al lenguaje audiovisual en general y al montaje en particular, tomando como partida los estudios señalados de Arielli y Manovich (2025), Gómez-Diago (2024) y Ortiz Arroba (2024), entre otros. Lo indudable es que una nueva estética, que se apoya en recursos de lo visual y en efectos de la capacidad generativa de los agentes de inteligencia artificial, revierte el tradicional sentido del montaje señalado por los teóricos del cine (Aumont, 2015; Eisenstein, 1991). Es verdad que, tras estas aproximaciones, es necesario fijar algunas de las conclusiones, expuestas a continuación.

Universidad Rey Juan Carlos
ASRI - Arte y Sociedad. Revista de Investigación
Núm. 31. Año 2026. ISSN: 2174-7563
CUVPAC Grupo de investigación CULTURA VISUAL & PRÁCTICAS ARTÍSTICAS CONTEMPORÁNEAS

Los ejemplos analizados tienen en común una preferencia por fórmulas de continuidad visual que generan fluidez, concebida desde la consecuencia de la unión de procesos de búsqueda y generación de versiones de *frames* a través de *prompts* de texto. La primera consecuencia consiste en una banda imagen que adquiere patrones de flujo intervenido o compuesto, propio de estéticas *maquínicas* de automatización y *randomización*, que afecta al montaje en su misma concepción. A tenor de lo visualizado en toda la muestra, pareciera que sirven de modelo tipos de transiciones que se oponen al montaje por corte (*cut*) como el *match-cut* (montaje por composición visual). Las piezas contienen un discurso ya no formado por unidades-plano de imagen en movimiento procedentes de tomas registradas con procedimientos fotoquímicos o digitales y cámaras convencionales, sino que es resultado de la unión de *frames* similares o de procesos donde se prima la generación de contenido visual equivalente. En paralelo, la unión de estos *frames* generan una especie de vibración visual o parpadeo, que puede relacionarse con el *stop motion* como técnica de animación exitosa en décadas anteriores. En este sentido, podría aventurarse que la animación expande su influencia como concepto estructurador de la generación de imágenes.

La inexistencia de la fase de registro en muchos de los clips de la muestra destapa que el montaje de estos vídeos resulta de la síntesis visual en forma tanto de transformación de *frames* como de la suma de versiones/iteraciones: el montaje no resulta de las decisiones de unión entre planos, sino que es consecuencia de las elecciones llegadas desde la automatización en la generación del propio material. Desaparece la lógica narrativa de causa-efecto, en un permanente fluir de juegos estéticos y visuales.

De manera general, estas piezas mantienen una exploración clásica de la materialidad de la imagen, que, aplicadas a las especificidades de la IA, anticipan quizás ciertas tendencias catalizadoras de una nueva siguiente fase, incipiente aún, de la cultura visual. Algunos rasgos destacados son la dimensión física de lo corpóreo, que se une a la búsqueda de enunciación conceptual de la visualidad. Por su parte, las imágenes fijas toman protagonismo, junto a un descenso de importancia de la combinación de planos, a lo que se une la posibilidad del reencuadre. Recursos como el *zoom* y el *travelling* ya no son planos diferentes, complementarios, sino consecuencia de la búsqueda automática por unir elementos dentro del espacio de la imagen, lo que influye notablemente en una nueva manera de concebir la escena, y permite indicar cierta influencia del universo de los videojuegos.

La automatización de técnicas que permite la IA profundiza en operaciones visuales que ya estaban presentes en la historia de la cultura visual. En cada género parece estar adaptando esta inclinación y consolidando ciertas fórmulas. En el caso del clip, su tendencia ya clásica a la experimentación extrema y asociativa —conceptual más que narrativa— recoge un terreno abonado para el trabajo con la postproducción, siguiendo la tendencia estética y narrativa de otros formatos como el videojuego. Se trasciende la importancia del código lingüístico y se vuelve a la simbología y la significación, además de la apuesta por la estética y la visualidad, apostando por la iconografía conceptual frente a la narrativa tradicional de combinación de unidades de significación narrativa.

Con todas estas técnicas, la IA parece que plantea soluciones de finalización basadas en el devenir, en el flujo. Estas operaciones resultan coherentes con una ecología de producción visual de la cantidad —de la proliferación— y de la variación y versión, que dominan la estética contemporánea desde hace décadas. En otras palabras, el videoclip musical realizado con herramientas de inteligencia artificial generativa

Universidad Rey Juan Carlos
ASRI - Arte y Sociedad. Revista de Investigación
Núm. 31. Año 2026. ISSN: 2174-7563
CUVPAC Grupo de investigación CULTURA VISUAL & PRÁCTICAS ARTÍSTICAS CONTEMPORÁNEAS

parece servir para desplazar la presencia del corte como unidad constitutiva y clave del montaje, que deja de ser proceso central y/o de culminación de estos textos culturales o, al menos, se camufla dentro de nuevos recursos expresivos.

En definitiva, la creación visual contemporánea con inteligencia artificial introduce conceptos inexplorados en torno a la autoría del material visual y la capacidad de agencia de los entes artificiales, pero lo hace siguiendo la línea de tradición procedente de la cultura de la red, donde la versión genera valor, a la vez que recoge dinámicas de viralidad e intertextualidad propias de la era de la postproducción, destacando el aprovechamiento de las posibilidades técnicas que reducen o enmascaran los procedimientos de montaje tradicional. En cualquier caso, siguiendo su tradición, el videoclip continúa ahondando en las lógicas de experimentación de los materiales visuales contemporáneos, y de comprensión del imaginario que acompaña a la música.

Referencias bibliográficas

Anantrasirichai, N. y Bull, D. (2022). Artificial Intelligence in the creative industries: a review. *Artificial Intelligence Review, 55*, 589–656. https://doi.org/10.1007/s10462-021-10039-7

Andrews, L. (2023). Unified AI creative environments: The rise of Superstudio. *Journal of Creative Tech Platforms, 8*(2), 41–48.

Arielli, E. y Manovich, L. (2025). *Artificial Aesthetics: Generative AI, Art and Visual Media*. Akal.

Aumont, J. (2015). *El montaje: la única invención del cine*. La marca editora.

Bourriaud, N. (2009). *Postproducción*. Adriana Hidalgo Editores.

Bruccoleri, A., Iacoviello, R., Messina, A., Metta, S., Montagnuolo, M. y Negro, F. (2022). *AI in Vision: High Quality Video Production & Content Automation*. AI4Media.

Caballero, J. (2023). Toward a New Dimension in Film Editing: Exploring the Possibilities of Artificial Intelligence. *Hipertext.net, 26*, 53–58. https://doi.org/10.31009/hipertext.net.2023.i26.08

Cazzaniga, M., Jaumotte, M. F., Li, L., Melina, M. G., Panton, A., Pizzinelli, C., Rockall, E. y Tavares, M. (2024). *Gen-AI: Artificial intelligence and the future of work*. International Monetary Fund.

Chen, B., Ziai, A., Tucker, R. y Xie, Y. (2022). Match Cutting: Finding Cuts with Smooth Visual Transitions. *arXiv.org*. https://doi.org/10.48550/arXiv.2210.05766

Crawford, K. (2022). *Atlas de inteligencia artificial: poder, política y costos planetario*. Fondo de Cultura Económica.

Deckker, D. y Sumanasekara, S. (2025). Dreams and Data: ghibli-style art, copyright, and the rise of viral AI Imagery. *International Journal of Global Economic Light, 11*(5), 54–65. https://doi.org/10.36713/epra0003

Eisenstein, S. (1991). *Hacia una teoría del montaje* (Vol. 2). Paidós.

Franganillo, J. (2022). Contenido generado por inteligencia artificial: oportunidades y amenazas. *Anuario ThinkEPI, 16*(1), 1–11. https://doi.org/10.3145/thinkepi.2022.e16a24

Gómez-Diago, G. (2024). Herramientas de inteligencia artificial para la creación, edición y traducción de vídeos. Una revisión. En S. Parrat Fernández, J. Mayoral Sánchez, y M. A. Chaparro Domínguez (Eds.), *Periodismo e inteligencia artificial. Aplicaciones y desafíos profesionales*. Comunicación Social Ediciones y Publicaciones. https://doi.org/10.52495/c4.emcs.25.p108

Jiang, Y., Li, X., Luo, H. y Yin, S. (2022). Quo Vadis Artificial Intelligence? Discover *Artificial Intelligence, 2*(4). https://doi.org/10.1007/s44163-022-00022-8

Jiménez-Jacinto, J., Seebach, S. y Martínez-Valverde, S. (2025). La Inteligencia Artificial Generativa y la redefinición del proceso creativo. *Espejo De Monografías De Comunicación Social,* (36), 141–152. https://doi.org/10.52495/c7.emcs.36.p114

Jironza Hidalgo, J. (2024). Análisis de la implementación de Inteligencia Artificial como herramienta de postproducción digital audiovisual. *Ñawi: arte diseño comunicación, 8*(2), 165–177. https://doi.org/10.37785/nw.v8n2.a9

Korsgaard, M. (2017). *Music Video After MTV. Audiovisual Studies, New Media, and Popular Music*. Routledge.

Kumar, R. (2023). AI-assisted video editing in content marketing platforms. *Visual Communication Today, 11*(3), 60–68.

Kumar, L., Goyal, P. y Kumar, R. (2020). Creativity in Machines: Music Composition Using Artificial Intelligence. *Asian Journal of Convergence in Technology, 6*(2), 36–40. https://doi.org/10.33130/ajct.2020v06i02.007

Lee, S. (2022). Scene detection technology in video post-production. *International Journal of Multimedia Tools, 30*(7), 1125–1138.

Lis Gutiérrez, J. y Pulido-Flórez, J. (2024). Creaciones artísticas musicales e Inteligencia Artificial (2019-2022). *El oído Pensante, 12*(1). https://doi.org/10.34096/oidopensante.v12n1.12176

López, M. (2024). Visual aesthetics in AI-generated video: Style control in platforms like Kaiber. *Digital Art & Intelligence Review, 10*(1), 66–72.

Manovich, L. (2018). *Ai Aesthetic*. Strelka Press.

Mazzi, F. (2023). Authorship in artificial intelligence-generated works: Exploring originality in text prompts and artificial intelligence outputs through philosophical foundations of copyright and collage protection. *The Journal of World Intellectual Property, 27*, 410–427. https://doi.org/10.1111/jwip.12310

Morgan, D. (2021). *The Lure of the Image: Epistemic Fantasies of the Moving Camera*. University of California Press.

Noesfm (2 de mayo de 2024). Sora de OpenAI transforma la producción musical en el último videoclip de Washed Out. *Noesfm*. https://noesfm.com/sora-de-openai-transforma-la-produccion-musical-en-el-ultimo-videoclip-de-washed-out/

Ortiz Arroba, F. (2024). La inteligencia artificial y la transformación de las expresiones artísticas. *Uru: Revista De Comunicación Y Cultura, 10*, 65–84. https://revistas.uasb.edu.ec/index.php/uru/article/view/4866

Sarzi-Ribeiro, R. (2023). Video performatividad. Estudios para videollamada de los artistas Lali Krotoszynski y Rodrigo Gontijo. *Revista Internacional de la Imagen, 8*(1), 39–48. https://doi.org/10.18848/2474-5197/CGP/v08i01/39-48

Schober, G. (2024). *Aesthetics of Digital Image Synthesis Conference* (2024, November 7-9). Klagenfurt, Austria.

Sedeño-Valdellós, A. (2025). *Manual de Realización audiovisual de videoclips musicales.* Editorial UOC.

Sedeño-Valdellós, A. Rodríguez López, J. y Roger Acuña, S. (2016). El videoclip postelevisivo actual. Propuesta metodológica y análisis estético. *Revista Latina de Comunicación Social, 71*, 332–348. https://doi.org/10.4185/RLCS-2016-1098

Smith, J. (2022). Automating video editing: Scene detection and segmentation. *Digital Video Quarterly, 9*(4), 28–35.

Top of the clips (2023). Skinny Pelembe Deadman Deadman Deadman Dylan Friese Greene. https://topoftheclips.com/skinny-pelembe-deadman-deadman-deadman-dylan-friese-greene/

Vásconez Merino, G., Moreira Sarango, K. y Estrada Falconí, A. (2025). La sensibilidad estética del anime como lenguaje cultural. El caso de Studio Ghibli. *Textual & Visual Media, 19*(2), 98–115. https://doi.org/10.56418/txt.19.2.2025

Vear, C. y Poltronieri, F. (2022). *The Language of Creative AI: Practices, Aesthetics and Structures*. Springer.

Zhao, X. y Ye, S. (2021). Space reconstruction of audiovisual media based on artificial intelligence and virtual reality. *Journal of Intelligent & Fuzzy Systems, 40*(4), 7285–7296. https://doi.org/10.3233/JIFS-189554

BIO



Ana Sedeño-Valdellos es Doctora en Comunicación Audiovisual y Profesora Titular en el Departamento de Comunicación Audiovisual y Publicidad de la Universidad de Málaga. Sus líneas de investigación tienen que ver con la música en relación con los medios audiovisuales, el vídeo musical y las prácticas audiovisuales en el panorama contemporáneo, con especial énfasis en elementos artísticos como la videoescena, la videodanza y las nuevas prácticas escénicas y audiovisuales. En relación con ellos ha publicado varios libros como *El lenguaje del vídeo musical*, *Manual de realización audiovisual de videoclips*, *Análisis del cine contemporáneo: estrategias estéticas, narrativa y puesta en escena* o *Historia del videoarte en España* y ha publicado en variadas revistas. Sus últimas líneas de investigación se han centrado en indagar en la capacidad de aplicación de la investigación artística a las ciencias de la comunicación y la exploración de nuevos formatos audiovisuales. Es coordinadora del Máster en Creación audiovisual y Artes Escénicas de la Facultad de Ciencias de la Comunicación de la Universidad de Málaga. Contacto: valdellos@uma.es.



Rafael Marfil-Carmona es Doctor en Didáctica de la Expresión Plástica (2015) y en Comunicación Audiovisual y Publicidad (2017) por la Universidad de Granada. Licenciado en Comunicación Audiovisual por la Universidad de Málaga y máster en Artes Visuales y Educación, es profesor titular en la Facultad de Ciencias de la Educación y ha sido subdirector del Departamento de Didáctica de la Expresión Musical, Plástica y Corporal hasta 2026. Miembro, entre otros colectivos, de la Sociedad para la Educación Artística (SEA), Grupo Comunicar, Icono 14 y la Red Interuniversitaria Euroamericana Alfamed, además de ser colaborador del grupo de investigación GICID, de la Universidad de Zaragoza y de SMEMIU, de UNED. En el ámbito profesional no docente ha trabajado, desde 1996, en prensa, televisión y comunicación corporativa. Su línea investigadora se centra en el análisis y la didáctica de las artes visuales y audiovisuales, así como la Educación Mediática y Artística, la fotografía, el cine, la publicidad, la Narrativa Audiovisual y la Cultura Digital. Es colaborador habitual de prensa en los medios de Granada, centrado en crítica cultural y opinión. Su web es imaginado.es. Contacto: rmarfil@ugr.es.