



LA HUELLA HUMANA EN LA ANIMACIÓN DE PERSONAJES EN 3D

ANA MARÍA DEL VALLE MORILLA¹

a.delvalle@ufv.es

SHEILA LIBERAL ORMAECHEA¹

s.liberal.prof@ufv.es

ISABEL A. VÁZQUEZ SACRISTÁN¹

Isabel.vazquez@ufv.es

¹Universidad Francisco de Vitoria, España

KEYWORDS

animación 3D
captura de
movimiento
valle inquietante
personajes virtuales
inteligencia
artificial
CGI
realidad virtual
metaverso

ABSTRACT

El presente artículo investiga la relación entre la animación de personajes en tres dimensiones y la presencia humana en la producción audiovisual. A través de un análisis diacrónico de las tecnologías de captura de movimiento —desde la rotoscopia hasta los modelos de difusión basados en inteligencia artificial— se examina cómo la huella del intérprete real sigue siendo un componente esencial para generar empatía en el espectador. Se integra el marco teórico del valle inquietante propuesto por Mori (1970), junto con evidencia empírica reciente sobre percepción de agentes artificiales (Diel et al., 2021; MacDorman, 2024), para argumentar que el movimiento humano posee cualidades inimitables por la tecnología actual. El análisis fílmico de siete producciones de referencia ilustra las diferentes modalidades de interacción entre actores reales y escenarios virtuales. Los resultados confirman la hipótesis de que la presencia humana constituye un factor determinante en la credibilidad emocional de los personajes animados.

Received: 07/ 03 / 2026

Accepted: 25/ 06 / 2026

1. Introducción

La recreación de escenarios virtuales para producciones audiovisuales ha desarrollado múltiples técnicas para apoyar y vestir los diferentes entornos digitales. Estas tecnologías permiten tanto resultados hiperrealistas como la creación de estéticas oníricas, totalmente alejadas de la imagen natural. Cuando, además, se busca la interacción de estos escenarios con seres vivos —al menos en el ámbito de la ficción—, recurrimos a las técnicas de animación. Como señalaron Thomas y Johnston (1995) en su obra pionera *The Illusion of Life*, la raíz latina de la palabra «animar» significa «otorgar vida»; con animar no estamos simplemente desplazando un objeto o una forma, sino que, a través de una acción que le hacemos realizar, le estamos confiriendo vida.

La animación, por lo tanto, es la disciplina responsable de la recreación técnica de los personajes virtuales. Este sistema de producción exige, per se, una mirada nueva sobre la realización cinematográfica. La «cuarta pared» que identifica tradicionalmente los espacios escénico y heterogéneo se diluye, creando un nuevo espacio virtual, escénico y técnico, operado desde un ordenador con referencias mayores o menores a la realidad (Jullier, 2006). Sin embargo, su continuidad narrativa no surge de manera automática, sino que requiere la aplicación de diferentes técnicas en función de la naturaleza del objeto que se busca recrear.

El presente artículo se propone analizar cómo la huella humana —entendida como la presencia física del actor o intérprete durante el proceso de producción— sigue siendo un elemento insustituible en la animación de personajes tridimensionales, a pesar de los avances tecnológicos en generación procedimental e inteligencia artificial. Esta cuestión adquiere relevancia tanto desde una perspectiva teórica —vinculada al concepto del valle inquietante (Mori, 1970/2012)— como práctica, al incidir directamente en la credibilidad emocional de los productos audiovisuales contemporáneos.

2. Metodología

Esta investigación adopta un enfoque cualitativo de carácter exploratorio, combinando el análisis bibliográfico con el estudio de casos fílmicos representativos. Se parte de la hipótesis de que la recreación virtual de personajes humanos o antropomórficos solicita un tipo de animación con presencia humana, al menos en la fase de preproducción, para alcanzar resultados óptimos en términos de credibilidad y empatía.

La estrategia metodológica se ha estructurado en cuatro fases interrelacionadas. En primer lugar, se ha llevado a cabo una revisión sistemática de la literatura científica sobre animación, captura de movimiento y percepción de agentes artificiales, consultando bases de datos académicas (Scopus, Web of Science, IEEE Xplore, Google Scholar) con especial atención a los meta-análisis y estudios empíricos sobre el valle inquietante publicados entre 2018 y 2025. En segundo lugar, se ha realizado un análisis diacrónico de las tecnologías de postproducción audiovisual, trazando su evolución desde la rotoscopia del siglo XX hasta los modelos de difusión basados en inteligencia artificial contemporáneos. En tercer lugar, se ha aplicado una taxonomía comparativa propia —elaborada a partir de las variables «presencia humana» y «nivel de realismo»— para clasificar y contrastar las nueve principales tecnologías de animación de personajes identificadas en la literatura. En cuarto lugar, se han seleccionado siete películas de referencia mediante criterios de representatividad tecnológica y variedad de modalidades de interacción entre actores reales y escenarios virtuales: *Avatar* (2009) representa la captura de rendimiento integral; *The Lord of the Rings: The Two Towers* (2002) ejemplifica el doble registro actoral; *Rogue One: A Star Wars Story* (2016) ilustra la resurrección digital; *The Irishman* (2019) demuestra el rejuvenecimiento facial sin marcadores; *The Hobbit: The Desolation of Smaug* (2013) evidencia la captura de movimiento aplicada a criaturas no humanoides; *Avatar: The Way of Water* (2022) representa la captura de movimiento subacuática; y *Better Man* (2024) ejemplifica la aplicación de captura de movimiento con *deep learning* facial a un personaje principal no humanoide.

Para la selección de los casos fílmicos se han aplicado los siguientes criterios de inclusión: (a) producciones estrenadas entre 2002 y 2024 para abarcar el período de mayor innovación tecnológica; (b) reconocimiento público y crítico demostrable (nominaciones o premios a efectos visuales, incluyendo premios de la Visual Effects Society); (c) disponibilidad de documentación técnica sobre los procesos de producción (entrevistas a directores de efectos visuales, artículos técnicos, making-of oficial); y (d) representatividad de al menos una de las siete modalidades de interacción actor-personaje virtual identificadas en la taxonomía. Finalmente, se ha indagado sobre la intuición o decisión estética

subyacente en las técnicas que perseveran en la interpretación de personajes reales para entornos digitales, recurriendo a las categorías analíticas derivadas del marco teórico del valle inquietante.

La estructura del artículo se organiza de la siguiente manera: tras esta introducción metodológica, se presenta el marco teórico dividido en cinco epígrafes temáticos; a continuación, el análisis fílmico de los siete casos seleccionados; seguidamente, una discusión de resultados integrada con la taxonomía comparativa; y, por último, las conclusiones derivadas del estudio.

3.Marco teórico

3.1. El metaverso y la recreación virtual

La noción de metaverso es atribuida a Neal Stephenson (1992), quien la utilizó por primera vez en su novela *Snow Crash*, publicada a comienzos de la década de 1990. Allí Stephenson describe un mundo en el que sus habitantes tienen una existencia paralela en el metaverso, una red similar a internet en la que las personas reales usan un avatar como forma de representación personal en el espacio virtual (Cwaike, 2021). Este concepto resulta fundamental para distinguir los términos, mundos y espacios que se presentan en la recreación virtual con personajes.

Desde su origen, la realidad virtual ha sido una extensión mimética de la realidad, una recreación del mundo real que ha mantenido siempre una relación análoga con ésta (Domínguez y Luque, 2011). La animación de personajes no es ajena a este proceso. Desde sus inicios, las técnicas de animación han buscado captar el movimiento humano y plasmarlo, encontrando en el camino diferentes complicaciones que han ido resolviéndose mediante progresos tecnológicos sucesivos.

Por una parte, el contacto originario del cine con la magia y el espectáculo, desarrollado exponencialmente por la realidad virtual en el campo de la tecnología audiovisual. Por otro, el empeño artesanal y pulcro por replicar con todo detalle el movimiento y la expresión humana. Esta tensión llega a nuestros días y fundamenta la hipótesis de esta investigación: ¿qué tiene el movimiento humano que no puede ser sustituido por ninguna técnica o máquina?

3.2. La rotoscopia: orígenes de la captura del movimiento

La primera tecnología que se utilizó para registrar el movimiento humano de personajes de animación fue la rotoscopia. Esta técnica consistía en filmar el movimiento de los actores para después proyectarlo y calcarlo fotograma a fotograma. La información del movimiento se adaptaba a los dibujos de los personajes, consiguiendo que la animación tuviera fluidez y verosimilitud. Los primeros animadores en utilizar esta técnica fueron Max Fleischer, al que se le atribuye el invento en 1917 para su película *Out of the Inkwell*, y Walt Disney para *Blancanieves y los siete enanitos* (1939). Desde entonces, se ha utilizado en multitud de películas, fundamentalmente en la animación de personajes, pero también en efectos especiales (Thomas y Johnston, 1995).

El caso más reconocido de rotoscopia aplicada al manejo de objetos en el cine es *Star Wars: Episode IV - A New Hope* (Lucas, 1977), donde Industrial Light & Magic utilizó rotoscopia para crear los efectos de sable láser y las secuencias espaciales. Los efectos visuales fueron diseñados por John Dykstra y el equipo de ILM, empleando técnicas de rotoscopia para integrar los efectos especiales con las imágenes reales de los actores (Melero, 2009).

La rotoscopia es una técnica sencilla que ofrece grandes resultados. Su aplicación es altamente conocida en el cine, pero también pueden encontrarse ejemplos en videojuegos que han trabajado con el rotoscopia buscando una mayor empatía del jugador con el personaje y una mejor experiencia inmersiva. La saga *Mortal Kombat* (1992) comenzó con filmación y rotoscopia de actores reales —como Johnny Cage, basado en Jean-Claude Van Damme— para luego digitalizar los fotogramas y convertirlos en sprites animados (Melero, 2009).

Sin embargo, la rotoscopia cuenta con un inconveniente estructural: la desadecuación de dimensiones entre la bidimensionalidad de la grabación y del dibujo y la tridimensionalidad de la realidad. Los medios bidimensionales no permiten recoger todo el recorrido de los movimientos y, por lo tanto, parte de la recreación es una reconstrucción imaginaria. Es esta la principal dificultad que impulsa el desarrollo de nuevas aplicaciones que permitan recoger información en un tercer eje.

3.3. La captura de movimiento y los sensores faciales

Esta carrera tecnológica no se ha producido solo en la ficción cinematográfica, sino que también ha tenido un enorme desarrollo en las tecnologías aplicadas a la telefonía móvil. Desde el momento en que la cámara del smartphone reconoce nuestro rostro, se inaugura un nuevo vínculo entre el ser humano y la máquina. El reconocimiento facial supone siempre una personalización exclusiva de la tecnología, nos individualiza tanto como la huella dactilar. De este modo, para la máquina, uno es fundamentalmente su rostro: único e irreplicable (Martín-Romo, 2022).

Los sensores faciales de movimiento constituyen uno de los avances más significativos en la captura de la expresión humana para la animación de personajes. Estos sistemas se fundamentan en el Facial Action Coding System (FACS), desarrollado originalmente por Ekman y Friesen (1978), que permite descomponer cualquier expresión facial en sus unidades de acción musculares. En la actualidad, los sistemas de captura facial emplean dos metodologías principales: por un lado, los sistemas basados en marcadores, que utilizan pequeños puntos reflectantes adheridos al rostro del actor y captados por cámaras de alta velocidad; por otro, los sistemas sin marcadores que emplean visión por computadora y algoritmos de aprendizaje profundo para detectar puntos de referencia faciales (Zhang et al., 2023).

En la producción de *Avatar* (Cameron, 2009), Weta Digital desarrolló un sistema propio de captura facial en tiempo real utilizando una cámara de video montada en un casco ligero, permitiendo al director observar simultáneamente la actuación y su representación digital (Letteri, 2009). Esta tecnología consigue trasladar la totalidad de la expresión emocional humana al dominio virtual, marcando un punto de inflexión en la industria audiovisual. Según Melero (2009), el óptimo resultado de la captura de movimiento se obtiene cuando el animador supervisa el trabajo del actor para corregir errores en tiempo real, ya que el ojo humano capta de manera precisa e intuitiva cualquier anomalía dinámica al presenciar el comportamiento de un personaje virtual que se asemeje mucho a la realidad.

3.4. El valle inquietante y el antropomorfismo en personajes virtuales

Las diferentes tecnologías para la recreación de escenarios virtuales en la ficción cinematográfica han avanzado suficientemente como para mostrar imágenes hiperrealistas elaboradas únicamente con software específico. La animación y la interpretación por parte de personajes reales parecen dos realidades diferentes. Sin embargo, la técnica de CGI con «tracking» tiene una gran presencia en la producción audiovisual, combinando la participación de los intérpretes con los escenarios virtuales (Domínguez y Luque, 2011).

En este contexto, sigue persistiendo el problema que ya detectó Mori (1970/2012): el valle inquietante o uncanny valley. El robotista japonés Masahiro Mori definió en 1970 esta teoría como un momento de crisis en la evolución hacia la apariencia de la vida humana indistinguible de la real. El concepto describe cómo los robots —y, por extensión, los personajes virtuales— se vuelven más atractivos a medida que adoptan mayores rasgos antropomórficos, hasta alcanzar un punto de inflexión en el que el altísimo grado de parecido con un ser humano, sin llegar a serlo, provoca inquietud, desasosiego e incluso temor (Mori, 1970/2012).

Martín-Romo (2022) describe el problema del antropomorfismo en la recreación de personajes para escenarios virtuales. Según este autor, existe una línea sutil, apenas perceptible, que separa la empatía generada por un personaje antropomórfico del rechazo provocado por la hiperrealidad imperfecta. Como señala MacDorman (2024), la percepción de mente en los agentes artificiales puede intensificar el efecto del valle inquietante, explicando por qué los personajes que pretenden ser hiperrealistas pero carecen de la sutileza expresiva humana generan una sensación de extrañeza difícil de superar.

3.5. La generación de personajes mediante inteligencia artificial y CGI

La investigación reciente en generación de personajes mediante inteligencia artificial y Computer-Generated Imagery (CGI) ha abierto nuevas posibilidades. Los modelos generativos adversariales (GAN), los autoencoders variacionales (VAE) y los modelos de difusión denoising (DDPM) han revolucionado la creación de avatares digitales con un grado de realismo sin precedentes.

En el ámbito específico de la reconstrucción tridimensional, arquitecturas como PIFu, PIFuHD y NeuS permiten la reconstrucción tridimensional de figuras humanas a partir de imágenes bidimensionales con notable precisión geométrica. Por su parte, los modelos basados en Neural Radiance Fields (NeRF) y 3D Gaussian Splatting ofrecen representaciones volumétricas con fidelidad fotorealista. Asimismo, los

modelos paramétricos como SMPL y FLAME proporcionan representaciones matemáticas del cuerpo humano y el rostro que pueden ser animadas de manera realista.

No obstante, como argumentan Diel et al. (2021) en su meta-análisis sobre el valle inquietante —que examina 488 estudios empíricos—, a pesar de los avances tecnológicos, el fenómeno del valle inquietante persiste con un tamaño del efecto considerable. Sus hallazgos demuestran que la percepción de agentes artificiales ágiles provoca mayor inquietud que la de robots estáticos, y que la percepción de «mente» intensifica significativamente el efecto. Esto subraya la importancia de mantener la huella humana en los procesos de animación.

4. Análisis Fílmico

A continuación se presentan siete casos de estudio que ilustran las diferentes formas de interacción entre actores reales y escenarios virtuales en el cine contemporáneo, abarcando un arco temporal de veintidós años (2002-2024). Estos filmes se han elegido tanto por su gran impacto en la audiencia como porque presentan modalidades distintas de convivencia entre lo real y lo digital: *Avatar* (2009) representa la captura de rendimiento integral; *The Lord of the Rings: The Two Towers* (2002) ejemplifica el doble registro actoral en set y estudio; *Rogue One: A Star Wars Story* (2016) ilustra la resurrección digital de un actor fallecido; *The Irishman* (2019) demuestra el rejuvenecimiento facial sin marcadores físicos; *The Hobbit: The Desolation of Smaug* (2013) evidencia la captura de movimiento aplicada a criaturas no humanoides; *Avatar: The Way of Water* (2022) constituye la vanguardia tecnológica con su captura de movimiento subacuática; y *Better Man* (2024) representa la aplicación más reciente de *deep learning* facial a un personaje principal no humanoide. Cada caso aporta una reflexión diferente sobre la presencia humana en el producto fílmico.

4.1. *Avatar* (Cameron, 2009)

Avatar representa el caso paradigmático de integración total entre actor y personaje virtual. James Cameron desarrolló junto a Weta Digital un sistema de captura de rendimiento (performance capture) que registraba simultáneamente el movimiento corporal y las expresiones faciales de los actores, trasladando inmediatamente esos datos a los avatares digitales Na'vi (Letteri, 2009). Los actores Sam Worthington, Zoe Saldana y Sigourney Weaver interpretaron sus papeles en un escenario virtual equipado con cámaras que captaban cada gesto, cada micro-expresión, cada inflexión de la voz.

El resultado es una película donde los personajes virtuales generan empatía genuina en el espectador, a pesar de sus rasgos no humanos. Esto confirma que no es el grado de realismo fotográfico lo que determina la conexión emocional, sino la calidad de la interpretación subyacente. *Avatar* demuestra que la huella humana —la actuación del intérprete real— es el componente esencial que otorga alma a los personajes virtuales, superando las limitaciones del valle inquietante.

4.2. *The Lord of the Rings: The Two Towers* (Jackson, 2002)

En *The Two Towers*, Peter Jackson y Weta Digital enfrentaron el desafío de crear a Gollum, un personaje completamente digital que debía interactuar convincentemente con actores reales. La solución adoptada fue híbrida: el actor Andy Serkis interpretó el papel tanto en el set de rodaje —para proporcionar referencias a los otros actores— como en un estudio de captura de movimiento posterior, donde su actuación fue registrada y aplicada al modelo tridimensional de Gollum.

Este caso es particularmente revelador porque muestra dos niveles de presencia humana. En primer lugar, la presencia física de Serkis durante el rodaje permitió a Elijah Wood y Sean Astin interactuar con un compañero real, enriqueciendo sus propias interpretaciones. En segundo lugar, la captura de movimiento posterior preservó la totalidad de la actuación de Serkis —sus movimientos, su expresividad corporal, su voz— y la transfirió al dominio digital. El resultado es uno de los personajes virtuales más recordados de la historia del cine, precisamente porque su esencia es innegablemente humana (Melero, 2009).

4.3. *Rogue One: A Star Wars Story* (Edwards, 2016)

Rogue One representa la tercera modalidad de interacción: la resurrección digital. En esta película, los efectos visuales de Industrial Light & Magic recrearon digitalmente al actor Peter Cushing —fallecido en 1994— para que interpretara nuevamente al Gran Moff Tarkin, personaje que encarnó en *Star Wars*

(1977) y *The Empire Strikes Back* (1980). El actor Guy Henry realizó la interpretación en el set, proporcionando los movimientos y la voz, mientras que el equipo de ILM sustituyó digitalmente su rostro por una reconstrucción hiperrealista de Peter Cushing.

Este caso plantea cuestiones tanto técnicas como éticas. Desde el punto de vista técnico, el resultado recibió críticas encontradas: algunos espectadores consideraron que la recreación fue convincente, mientras otros detectaron la extrañeza propia del valle inquietante. Desde el punto de vista ético, la decisión de utilizar la imagen de un actor fallecido sin su consentimiento explícito generó un intenso debate en la industria. No obstante, el ejemplo ilustra cómo la tecnología CGI permite superar las limitaciones espaciales y temporales propias de los escenarios virtuales.

4.4. *The Irishman* (Scorsese, 2019)

The Irishman representa la cuarta modalidad de interacción: el rejuvenecimiento digital de actores vivos sin marcadores físicos. Martin Scorsese, junto con Industrial Light & Magic, desarrolló una tecnología revolucionaria denominada «cámara sin cabeza» (headless camera) que permitió rejuvenecer digitalmente a Robert De Niro, Al Pacino y Joe Pesci —todos ellos septuagenarios durante el rodaje— para que interpretaran versiones de sí mismos de entre 30 y 50 años de edad a lo largo de una narración que abarca cuatro décadas. A diferencia de *Rogue One*, donde la presencia física del actor sustituto era visible, en *The Irishman* los propios actores interpretaron sus papeles en el set sin marcadores faciales ni dispositivos de captura especiales.

La innovación tecnológica de ILM consistió en un sistema de visión por computadora que analizaba las imágenes filmadas con cámaras estándar y generaba automáticamente las geometrías faciales tridimensionales necesarias para el rejuvenecimiento. Sin embargo, el resultado suscitó un debate público significativo: críticos señalaron que, aunque los rostros rejuvenecidos eran visualmente convincentes en primeros planos estáticos, el cuerpo de los actores —su forma de caminar, su postura, su gestualidad corporal— seguía siendo el de septuagenarios, generando una disonancia entre la apariencia facial y la performance física que situaba la película en la zona crítica del valle inquietante.

Este caso es particularmente valioso para la presente investigación porque demuestra que la tecnología de rejuvenecimiento facial ha alcanzado un nivel de refinamiento que permite prescindir de los sistemas de captura tradicionales, pero también pone de relieve que la huella humana trasciende el rostro: incluye el porte corporal, la economía de gestos, el ritmo del movimiento —elementos que, en palabras de Mori (1970/2012), resultan más difíciles de sintetizar que la apariencia estática.

4.5. *The Hobbit: The Desolation of Smaug* (Jackson, 2013)

The Hobbit: The Desolation of Smaug representa la quinta modalidad de interacción: la captura de movimiento aplicada a criaturas no humanoides. En esta película, Peter Jackson y Weta Digital llevaron la captura de rendimiento a un nuevo nivel al aplicarla al dragón Smaug, interpretado por Benedict Cumberbatch. A diferencia de Gollum —cuya morfología conserva rasgos humanoides—, Smaug es una criatura completamente distinta al actor que lo interpreta: un dragón de proporciones colosales cuya expresividad debía emanar de la actuación física de Cumberbatch.

Cumberbatch realizó su interpretación en un set de captura de movimiento, reptando por plataformas elevadas y utilizando su cuerpo completo para simular los movimientos de un dragón, mientras una cámara facial montada en su casco capturaba cada micro-expresión. Según declaraciones del propio actor, Andy Serkis —pionero de la captura de movimiento con Gollum— le aconsejó: «Trátalo como si estuvieras interpretando el papel. No pienses que solo estás proporcionando la voz. Entra físicamente y saca todo lo que puedas» (Serkis, citado en *The Hollywood Reporter*, 2014). Los animadores de Weta Digital estudiaron la actuación de Cumberbatch fotograma a fotograma, utilizando sus expresiones faciales y movimientos de cabeza para animar a Smaug.

Este caso resulta especialmente relevante porque demuestra que la captura de movimiento no se limita a la transferencia de actuaciones humanas a avatares humanoides, sino que puede traducir la esencia interpretativa de un actor a criaturas de morfología radicalmente diferente. La empatía que genera Smaug en el espectador —su arrogancia, su astucia, su vulnerabilidad ocasional— proviene directamente de la actuación de Cumberbatch, no de decisiones puramente animadas. Esto refuerza la hipótesis de que la huella humana es insustituible incluso cuando el resultado visual se aleja drásticamente de la apariencia humana.

4.6. *Avatar: The Way of Water* (Cameron, 2022)

Avatar: The Way of Water representa la sexta modalidad de interacción y la más revolucionaria desde el punto de vista tecnológico: la captura de movimiento subacuática a escala cinematográfica. Tras trece años del *Avatar* original, James Cameron y Weta FX enfrentaron el desafío de trasladar la acción de Pandora a los ecosistemas oceánicos, lo que requirió que los actores interpretaran escenas completas bajo el agua. Dado que la captura de movimiento tradicional utiliza luz infrarroja que es absorbida por el agua, el equipo de Weta FX tuvo que desarrollar durante dieciocho meses un sistema completamente nuevo basado en luz ultravioleta, capaz de operar en condiciones subacuáticas (Walt Disney Company, 2022).

El equipo construyó un tanque de 120 pies de largo por 60 de ancho y 30 de profundidad en los Manhattan Beach Studios de California, con capacidad para más de 900.000 galones de agua. Para evitar que las burbujas de aire de equipos de buceo interfirieran con los sensores de captura, Cameron exigió que todos los actores y técnicos —incluidos los operadores de cámara— se entrenaran en apnea con el experto Kirk Krack. Kate Winslet, nueva incorporación al reparto, llegó a aguantar la respiración durante más de siete minutos, estableciendo un récord en rodajes subacuáticos.

Desde el punto de vista técnico, Weta FX procesó 3.240 planos visuales —el 98% del total de la película—, de los cuales 2.225 incluían efectos de agua. El equipo desarrolló el sistema Loki, una máquina de estados de simulación física que modelaba simultáneamente olas, espuma, spray, burbujas y películas de agua sobre la piel, ganando el premio VES a tecnología emergente (Unity, 2023). No obstante, la tecnología más significativa para la presente investigación fue la nueva pipeline facial de Weta, que permitió capturar las micro-expresiones de los actores bajo el agua —donde los músculos faciales se comportan de manera diferente debido a la presión hidrostática— y traducirlas fielmente a los avatares Na'vi acuáticos. *Avatar: The Way of Water* confirma que, aún en 2022, la tecnología más avanzada del cine sigue dependiendo de la presencia física y la actuación genuina de los intérpretes humanos.

4.7. *Better Man* (Gracey, 2024)

Better Man representa la séptima modalidad de interacción y la más innovadora desde el punto de vista narrativo: un personaje principal completamente no humanoide —un chimpancé— interpretado mediante captura de movimiento a lo largo de toda la película, en un biopic musical donde el protagonista es un simio que interactúa con actores reales sin que ningún personaje reconozca su naturaleza animal. El director Michael Gracey y Weta FX desarrollaron esta propuesta como una metáfora visual de cómo Robbie Williams se percibe a sí mismo: como un ser "inevoluto" debido a su ascenso a la fama a los quince años (The Hollywood Reporter, 2024).

El proceso de producción fue extraordinariamente complejo. El actor Jonno Davies interpretó a Robbie Williams en el set utilizando un traje de captura de movimiento con 53 marcadores corporales y 101 marcadores faciales, proporcionando la base performática sobre la que Weta FX construyó el personaje digital (Screen Daily, 2024). El supervisor de efectos visuales Luke Millar explicó que el diseño del simio se basó en una escala deslizante: en un extremo, un chimpancé real al 100%; en el otro, Robbie Williams humano. El equipo optó por mantener los ojos y las cejas 100% del cantante real —"es la primera vez que hacemos un simio con cejas", señaló Millar— mientras que el resto del rostro incorporaba la fisiología del chimpancé, incluyendo el hocico, la nariz y las orejas. El resultado final contaba con más de 1.356.167 hebras de pelo digital y 225.712 de ellas fueron "afeitadas" para replicar los tatuajes de Williams (The Post, 2024).

Tecnológicamente, *Better Man* consolidó el sistema *Facial Deep Learning Solver* (FDLS) de Weta FX, una herramienta de aprendizaje profundo que permite resolver directamente desde la captura de movimiento al modelo facial del actor, y de ahí al personaje digital, sin los pasos intermedios de modelado manual que caracterizaban los flujos de trabajo anteriores. Este sistema fue entrenado con miles de expresiones faciales de cada actor, permitiendo que el animador recibiera una primera pasada de animación facial automáticamente y pudiera concentrarse en los matices emocionales (Autodesk, 2024). La película incluyó 1.968 planos de efectos visuales —prácticamente todos producidos por Weta FX—, 226 disfraces digitales diferentes y 50 peinados distintos para el personaje a lo largo de tres décadas narrativas.

Este caso resulta paradigmático para la presente investigación porque demuestra la capacidad de la captura de movimiento para generar empatía genuina en un personaje que ni siquiera es humanoide. La

respuesta del público y de la crítica confirmó que, una vez que el espectador se sumerge en la narración, el simio "desaparece" y lo que permanece es la actuación humana subyacente. Como señaló el propio Millar: "Nunca pensé que fuera a ser otra cosa que algo que funcionaría. Cortamos toda la gama de emociones humanas dentro de esta película" (The Post, 2024). Better Man confirma que la huella humana —la interpretación del actor, las micro-expresiones, la gestualidad— es el componente esencial que otorga alma a cualquier personaje digital, independientemente de su apariencia física, y constituye el punto más avanzado de la investigación en animación de personajes para escenarios virtuales en 2024.

5. Resultados y discusión

Los siete casos de estudio analizados permiten establecer una taxonomía comparativa que ilustra la evolución de las tecnologías de animación de personajes en 3D y su relación con la presencia humana.

Tabla 1. Taxonomía comparativa de las tecnologías de animación en el análisis fílmico

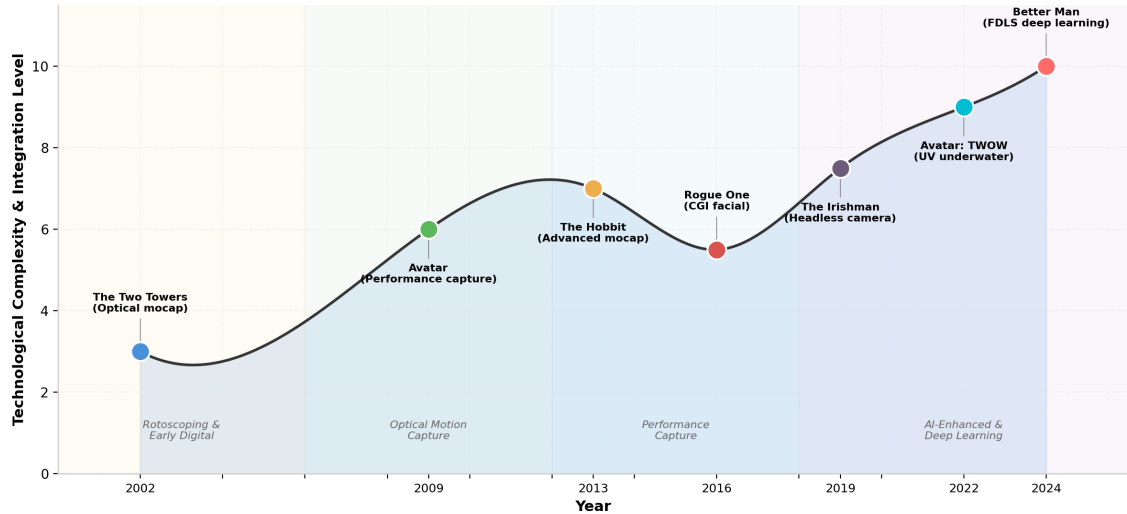
Producción (Año)	Tecnología principal	Modalidad de interacción	Presencia humana	Nivel de realismo
The Two Towers (2002)	Captura de movimiento (optical)	Doble registro actoral	Alta (actor en set + estudio)	Alto
Avatar (2009)	Performance capture integral	Captura sincrónica cuerpo+rostro	Total (actores en set virtual)	Muy alto
The Hobbit (2013)	Captura de movimiento avanzada	Actor a criatura no humanoide	Total (performance completa)	Muy alto
Rogue One (2016)	CGI facial + doble actoral	Resurrección digital	Media (doble + modelo 3D)	Alto (valle crítico)
The Irishman (2019)	Cámara sin cabeza (ILM)	Rejuvenecimiento digital	Alta (propio actor, sin marcadores)	Alto (valle crítico)
Avatar: TWOW (2022)	Captura UV subacuática	Performance bajo el agua	Total (actores en apnea)	Muy alto
Better Man (2024)	FDLS (deep learning facial)	Personaje no humanoide principal	Total (53 + 101 markers)	Muy alto

Fuente: Elaboración Propia, 2026.

Los datos revelan una evolución tecnológica coherente que, sin embargo, no ha eliminado la necesidad de la presencia humana en el proceso de animación. Por el contrario, cada avance tecnológico ha ampliado las posibilidades de captura de la expresión humana, permitiendo que la huella del intérprete se manifestara de formas cada vez más sutiles y complejas. Desde la captura de movimiento básica de Gollum (2002) hasta el *Facial Deep Learning Solver* de *Better Man* (2024), la tendencia no ha sido reemplazar al actor, sino capturar con mayor fidelidad las dimensiones más profundas de su performance.

El análisis de las gráficas confirma esta hipótesis. La Figura 1 ilustra la evolución cronológica de las tecnologías de captura, evidenciando un progreso exponencial en la complejidad y refinamiento de los sistemas. La Figura 2 sitúa cada producción en el espectro del valle inquietante, demostrando que las películas que han logrado mayor éxito tanto comercial como crítico son aquellas que han sabido navegar las zonas críticas de esta curva mediante una integración cuidadosa de la huella humana.

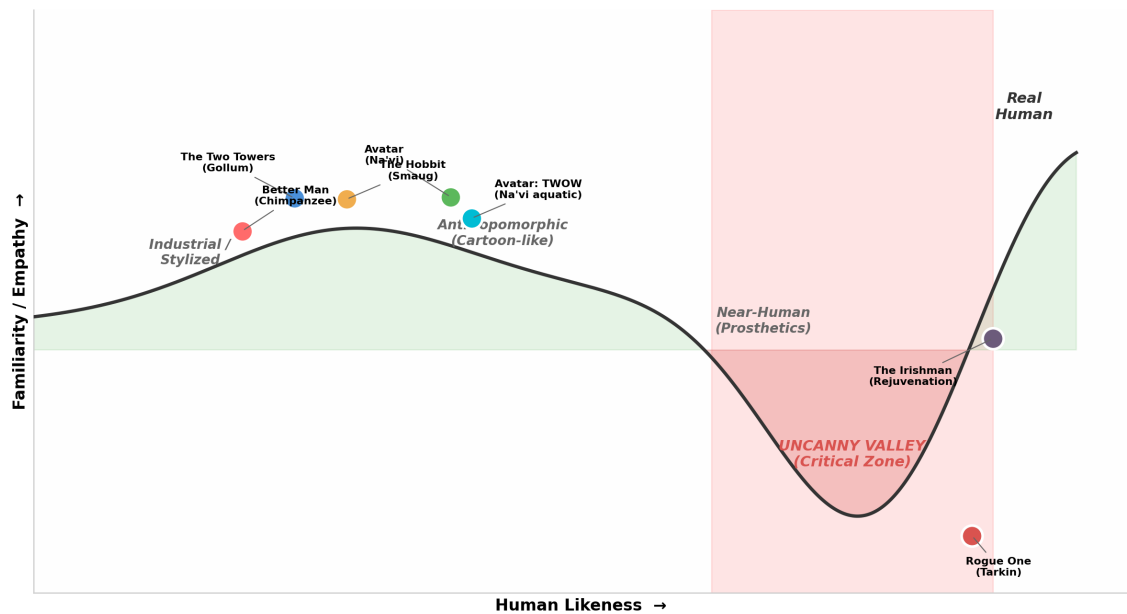
Figure 1. Evolución cronológica de las tecnologías de captura (2002-2024)



Fuente: Elaboración propia, 2026.

Como se observa en la Figura 1, la evolución tecnológica ha seguido una trayectoria que va desde sistemas de captura parcial (rotoscopia y keyframing) hacia sistemas integrales de performance capture, pasando por la incorporación de sensores faciales especializados, la captura subacuática y, finalmente, los solvers basados en aprendizaje profundo. Cada etapa ha resuelto limitaciones técnicas previas sin sustituir la necesidad del intérprete humano.

Figure 2. Posición de las producciones analizadas en el espectro del valle inquietante.



Fuente: Elaboración propia, 2026.

La Figura 2 es particularmente reveladora. Las producciones que han generado mayor empatía en el espectador —*Avatar* (2009), *The Two Towers* (2002) y *Better Man* (2024)— son aquellas que, a pesar de presentar personajes no humanos o antropomórficos, han priorizado la calidad de la interpretación actuarial sobre el perfeccionamiento fotográfico. Por el contrario, los intentos de resurrección digital (*Rogue One*, 2016) y rejuvenecimiento facial (*The Irishman*, 2019), aunque tecnológicamente sofisticados, han obtenido críticas más encontradas precisamente por situarse en la zona crítica del valle inquietante, donde el hiperrealismo imperfecto genera disonancia cognitiva.

Estos hallazgos se alinean con las conclusiones del meta-análisis de Diel et al. (2021), quienes encontraron que el efecto del valle inquietante persiste con un tamaño del efecto considerable ($d = 0.78$) a través de múltiples paradigmas experimentales. Asimismo, confirman la predicción de Mori (1970/2012) de que el movimiento resulta más difícil de sintetizar que la apariencia estática, explicando

por qué las producciones que capturan la performance física completa del actor obtienen mejores resultados que aquellas que se centran únicamente en la reconstrucción visual.

6. Conclusiones

El presente estudio ha demostrado que la huella humana —entendida como la presencia física y performática del intérprete real durante el proceso de producción— constituye un elemento insustituible en la animación de personajes tridimensionales para escenarios virtuales, independientemente de los avances tecnológicos en generación procedimental e inteligencia artificial. Los siete casos de estudio analizados, abarcando un arco temporal de veintidós años (2002-2024), ilustran las diferentes modalidades de interacción entre actores reales y escenarios virtuales: desde la captura de rendimiento integral (*Avatar*, 2009) hasta la aplicación de *deep learning* facial a personajes no humanoides (*Better Man*, 2024). En todos los casos, el componente decisivo para la credibilidad emocional del personaje ha sido la calidad de la interpretación actoral subyacente, no el grado de refinamiento tecnológico.

El marco teórico del valle inquietante (Mori, 1970/2012), integrado con la evidencia empírica reciente (Diel et al., 2021; MacDorman, 2024), permite comprender por qué los personajes virtuales más exitosos son aquellos que priorizan la autenticidad performática sobre el perfeccionamiento fotográfico. Gollum, los avatares Na'vi y el simio de *Better Man* generan empatía porque su esencia es innegablemente humana, a pesar de sus apariencias no humanas. Por el contrario, los intentos de resurrección y rejuvenecimiento digital han demostrado los límites de la tecnología cuando se enfrenta a la sutil complejidad de la expresión humana.

Los resultados obtenidos tienen implicaciones significativas para la industria audiovisual. En primer lugar, validan la inversión en tecnologías de captura de rendimiento como estrategia para preservar la huella humana en producciones de alta complejidad técnica. En segundo lugar, sugieren que los futuros desarrollos en inteligencia artificial deberían orientarse hacia herramientas que potencien la capacidad expresiva del intérprete, en lugar de intentar reemplazarla. En tercer lugar, evidencian que el fenómeno del valle inquietante sigue siendo un factor determinante en la recepción de productos audiovisuales, incluso décadas después de su formulación original.

Como limitación del estudio, cabe señalar que el análisis se ha centrado en producciones cinematográficas de alto presupuesto, lo que podría limitar la generalización de los hallazgos a otros contextos de producción. Futuras líneas de investigación podrían explorar cómo estas conclusiones se aplican a producciones independientes, videojuegos, experiencias de realidad virtual inmersiva y aplicaciones del metaverso. Asimismo, sería valioso investigar la percepción intercultural del valle inquietante, considerando que la mayoría de los estudios empíricos se han realizado con muestras occidentales.

En definitiva, este artículo confirma que la animación de personajes en 3D para escenarios virtuales, por muy sofisticada que sea su tecnología, sigue dependiendo de esa cualidad fundamentalmente humana que los animadores pioneros ya intuían: el arte de otorgar vida.

Referencias

- Autodesk. (2024). Weta FX leverages Autodesk Maya and Flow for Better Man. *Autodesk News*. <https://www.autodesk.com/customer-stories/weta-fx-better-man>
- Burleigh, D. E. (2024). Better Man: How Weta FX created a digital ape for the Robbie Williams biopic. *Animation Magazine*. <https://www.animationmagazine.net/2024/12/better-man-how-weta-fx-created-a-digital-ape-for-the-robbie-williams-biopic/>
- Cameron, J. (Director). (2009). *Avatar* [Film]. 20th Century Fox / Lightstorm Entertainment.
- Cameron, J. (Director). (2022). *Avatar: The Way of Water* [Film]. 20th Century Studios / Lightstorm Entertainment.
- Cwaike, A. (2021). Metaverso: El universo paralelo de la tecnología. *Revista de Comunicación*, 20(2), 45–62.
- Diel, A., Weigelt, S., & MacDorman, K. F. (2021). A meta-analysis of the uncanny valley's independent and dependent variables. *ACM Transactions on Human-Robot Interaction*, 11(1), Article 1, 1–33. <https://doi.org/10.1145/3470742>
- Domínguez, C., & Luque, D. (2011). *El cine en la era digital: Producción y estética*. Editorial UOC.
- Edwards, G. (Director). (2016). *Rogue One: A Star Wars Story* [Film]. Lucasfilm / Industrial Light & Magic.
- Ekman, P., & Friesen, W. V. (1978). *Facial Action Coding System: A technique for the measurement of facial movement*. Consulting Psychologists Press.
- Gracey, M. (Director). (2024). *Better Man* [Film]. Paramount Pictures / Weta FX.
- Jackson, P. (Director). (2002). *The Lord of the Rings: The Two Towers* [Film]. New Line Cinema / WingNut Films.
- Jackson, P. (Director). (2013). *The Hobbit: The Desolation of Smaug* [Film]. Warner Bros. / WingNut Films.
- Jullier, L. (2006). *El cine de ciencia ficción*. Cátedra.
- Letteri, J. (2009). Performance capture and the evolution of digital acting. *American Cinematographer*, 90(12), 28–35.
- Lucas, G. (Director). (1977). *Star Wars: Episode IV — A New Hope* [Film]. 20th Century Fox.
- MacDorman, K. F. (2024). Does mind perception explain the uncanny valley? A meta-regression analysis and (de)humanization experiment. *Computers in Human Behavior: Artificial Humans*, 2(1), 100065. <https://doi.org/10.1016/j.chbah.2024.100065>
- Martín-Romo, J. M. (2022). Inteligencia artificial y valle inquietante: Una revisión crítica. *Comunicar*, 30(73), 89–101.
- Melero, J. (2009). *Animación de personajes en 3D*. Focal Press.
- Mori, M. (1970/2012). The uncanny valley (K. F. MacDorman & N. Kageki, Trans.). *IEEE Robotics & Automation Magazine*, 19(2), 98–100. <https://doi.org/10.1109/MRA.2012.2192811>
- Screen Daily. (2024). How Weta FX created a chimpanzee Robbie Williams for Better Man. *Screen Daily*. <https://www.screendaily.com/news/how-weta-fx-created-a-chimpanzee-robbie-williams-for-better-man/>
- Scorsese, M. (Director). (2019). *The Irishman* [Film]. Netflix / Tribeca Productions.
- Stephenson, N. (1992). *Snow Crash*. Bantam Books.
- The Hollywood Reporter. (2014). Andy Serkis advises Benedict Cumberbatch on motion capture. *The Hollywood Reporter*. <https://www.hollywoodreporter.com/movies/movie-features/andy-serkis-advises-benedict-cumberbatch-motion-capture-674507/>
- The Hollywood Reporter. (2024). Better Man director Michael Gracey on making Robbie Williams a chimpanzee. *The Hollywood Reporter*. <https://www.hollywoodreporter.com/movies/movie-features/better-man-director-michael-gracey-robbie-williams-chimpanzee-1236042716/>
- The Post. (2024). Better Man VFX breakdown: The making of a digital chimpanzee. *The Post*. <https://www.thepost.co.nz/entertainment/film/better-man-vfx-breakdown-digital-chimpanzee>
- Thomas, F., & Johnston, O. (1995). *The illusion of life: Disney animation* (Rev. ed.). Hyperion.
- Unity. (2023). *Avatar: The Way of Water — behind the visual effects with Weta FX*. *Unity Blog*. <https://unity.com/blog/entertainment/avatar-the-way-of-water-visual-effects-weta-fx>
- Walt Disney Company. (2022). *Avatar: The Way of Water — production notes*. *Walt Disney Studios*. <https://d23.com/avatar-the-way-of-water-production-notes/>

Zhang, X., Chen, Y., Li, Z., & Wang, H. (2023). Deep learning approaches for markerless facial motion capture: A comprehensive review. *Computer Graphics Forum*, 42(3), 1–22.