

MÁSTER EN COMPUTACIÓN GRÁFICA Y SIMULACIÓN

2018

Trabajo de Final de Máster

Investigación, evaluación e implementación de
métodos que simulen seis grados de libertad
en fotos y vídeo para Realidad Virtual

Autor: Gregorio Iniesta Ovejero

Tutor: Diego Bezares Sánchez

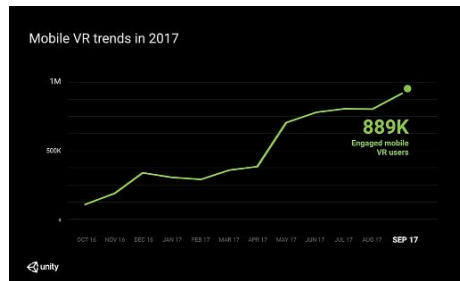
Índice general

1. Resumen	1
2. Introducción	3
2.1. Grados de libertad	3
2.2. Foto y vídeo en realidad virtual	4
2.3. Qué aporta este proyecto	5
3. Planteamiento del problema	7
4. Objetivos	9
5. Desarrollo	11
6. Resultados	13
7. Conclusiones	15
Bibliografía	17

1. Resumen

2. Introducción

La realidad virtual es una tecnología que desde hace unos años ha estado intentando hacerse un hueco en la industria del entretenimiento. Unity expuso estadísticas de consumo de realidad virtual durante el pasado año, proporcionando cifras como un 134 % de crecimiento en sobremesa y consola y un 295 % de crecimiento en las instalaciones de realidad virtual en móvil lo que hace casi un millón de usuarios a finales de 2017.



2.1. Grados de libertad

Un factor clave a tener en cuenta cuando hablamos de realidad virtual son los grados de libertad, que definen la capacidad de movimiento a la hora de interactuar. Generalmente se habla de tres y seis grados de libertad. En el caso de tres grados de libertad, hace referencia a los giros sobre el eje principal (viraje o *yaw*, inclinación o *pitch* y cabeceo o *roll*), mientras que cuando se amplía a seis grados de libertad hace referencia al desplazamiento en los tres ejes. Esto aplicado a las gafas de realidad virtual informa de la capacidad del dispositivo de captar los giros de la cabeza

en todas las direcciones (tres grados de libertad) o si además capta el desplazamiento por la sala (seis grados de libertad).

El mercado de dispositivos de realidad virtual se está enfocando hacia seis grados de libertad. La alta gama ya dispone de ellos desde el principio mientras la gama media y baja ya está adaptándose como demuestra Oculus Santa Cruz o Vive Focus.

2.2. Foto y vídeo en realidad virtual

El contenido que se genera todavía está basado en gran parte en técnicas ya conocidas como reproducción de vídeo y fotos cuya máxima adaptación consiste simplemente en poner una imagen ligeramente diferente en cada ojo.

La característica principal de este tipo de contenido es que cada imagen está tomada desde un punto fijo en el espacio. Esto provoca una problemática que consiste en que el usuario únicamente tiene tres grados de libertad a la hora de visualizarlo en unas gafas de realidad virtual. Además de esta restricción debido a las técnicas de grabación que se usan, el cabeceo o *roll* provoca ver imágenes duplicadas y puede provocar incomodidad o incluso mareo.

Debido a las estadísticas de consumo del contenido multimedia grandes empresas como Google, Facebook y Disney entre otras, están trabajando en mejorar los sistemas de visualización de vídeo y fotos mediante vídeo volumétrico, lightfields o fotogrametría.

La mayoría de proyectos que están disponibles actualmente utilizan técnicas que transforman cada frame en polígonos y desplazan los vértices en función de la profundidad que le indica un mapa de profundidad. Esto provoca en demasiadas ocasiones una deformación de elementos pequeños y en algunas ocasiones se ocultan elementos que deberían ser visibles de otro modo.

2.3. Qué aporta este proyecto

Este trabajo toma la demo “Welcome to lightfields” como referencia pero creando un proyecto de código libre. Trata de conseguir una sensación real de tres dimensiones habilitando seis grados de libertad en un espacio reducido de desplazamiento a partir de un vídeo estereoscópico plano en dos dimensiones y su mapa de profundidad. Para ello se hará un paralaje en tres dimensiones píxel a píxel en función del mapa de profundidad y la posición del usuario.

Durante el desarrollo del proyecto se llevan a cabo pruebas con diferentes técnicas y se evalúa la viabilidad en diferentes dispositivos (plataformas móviles y de sobremesa), el realismo del resultado así como la escalabilidad de los métodos.

3. Planteamiento del problema

La fotografía y el vídeo para realidad virtual actualmente tienen muchas limitaciones tanto para producirlo como para visualizarlo. Una de estas limitaciones y del cual este proyecto se ocupa es la falta de libertad movimiento.

La foto y el vídeo sólo tienen tres grados de libertad y de ellos cabeceo o *roll* (inclinación la cabeza sobre los hombros) no funciona como cabría esperar y puede provocar desde mareos hasta ver las imágenes duplicadas. La captura de imágenes tanto reales como virtuales se hace con cámaras que irremediablemente están en un punto concreto del espacio y eso en principio limita el movimiento del usuario.

Por ello, este trabajo se centra en mejorar la visualización del vídeo con un sistema que permita al usuario desplazarse físicamente dentro de un área reaccionando el vídeo a ese posicionamiento en tiempo real.

4. Objetivos

Este proyecto ha sido creado con el objetivo de investigar, evaluar e implementar software para proporcionar seis grados de libertad en la visualización de contenidos a partir de imágenes planas y su mapa de profundidad aplicado en tecnologías de realidad virtual.

5. Estado del Arte

La realidad virtual actualmente abarca muchas areas y se puede interactuar con ella de muchas formas. La forma más común y en la cual se va a centrar este documento es en la representación de imágenes mediante un casco o gafas de realidad virtual.

5.1. Mapas de profundidad

Debido a la cantidad de técnicas que utilizan mapas de profundidad es interesante explicar en que consisten.

Los mapas de profundidad son imagenes que en cada pixel se encuentra codificada la profundidad de la foto en ese punto. Generalmente se utiliza una escala de grises o de rojos aunque se pueden recurrir a métodos más complejos. [2]

En el caso de la escala de grises (??), los tonos más oscuros representan elementos en el fondo de la imagen, mientras que los tonos mas claro representan elementos más cercanos.

En el caso de una imagen generada por ordenador, es fácil obetener un buen mapa de profundidad. Sin embargo en el caso de las imágenes captadas por cámaras reales, existe la posibilidad de que la cámara esté preparada y también obtenga la profundidad o en caso contrario habría que aplicar algoritmos que calculen la profundidad en cada píxel. Estos algoritmos no son perfectos y en ocasiones el mapa obtenido tiene ruido y puede afectar a las técnicas que lo usen.

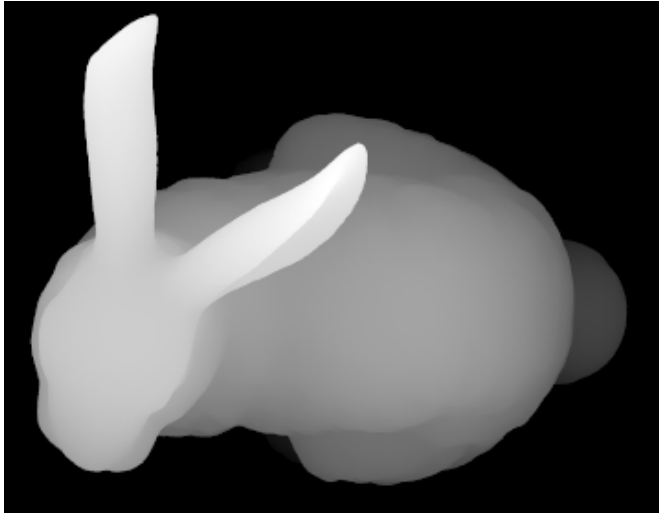


Figura 5.1: Ejemplo de mapa de profundidad cenital

5.2. Project Sidewinder

Adobe presento en 2017 [1] una demo que utilizaba un depthmap de manera muy sencilla para permitir seis grados de libertad dentro de un video 360°. No proporcionan mucha información ya que es una prueba de concepto.

El desplazamiento punto a punto parece correcto pero se ve una distorsión en los bordes probablemente debido al estado temprano del proyecto.

5.3. Nube de puntos

[5]Esta técnica consiste en crear un punto en el espacio por cada pixel de la foto o el vídeo y utilizar la imagen y su mapa de profundidad para

seleccionar el color y el desplazamiento del punto en el espacio.

Esta implementación por contra provoca que aparezcan muchos huecos, generalmente se suele acompañar la implementación con una selección del tamaño del punto como muestra [2] para ver una imagen más sólida.

Otro de los problemas que tiene esta técnica es la cantidad de puntos que deben ser tratados, ya que una resolución QHD requiere cuatro millones de puntos y la resolución recomendada actualmente es 4K.

5.4. Seis Grados de libertad con Fotogrametría

Una de las formas de conseguir seis grados de libertad es reconstruir la escena mediante fotogrametría como muestra Facebook en [4] para procesar la imagen y obtener una malla que pueda ser mostrada usando técnicas comunes.

Este procedimiento utiliza una técnica nueva que han llamado mapa de profundidad de límite inferior que la profundidad de cada pixel debe ser estrictamente mayor que en el depthmap normal. Crear este mapa de profundidad le ayuda al ensamblado de las imágenes para crear la imagen final 360°. Además mezclando el mapa de profundidad con otros algoritmos son capaces de obtener un mapa de normales bastante preciso.

Finalmente crean la malla a partir de una nube de puntos. Esta malla como veremos en otras técnicas, tiene agujeros detrás de los objetos que no pueden ser rellenados por falta de información. En este caso optan por

difuminar de manera sutil las zonas desconocidas dando un buen resultado.

Todo esto lo aprovechan para poder generar un entorno tridimensional con el que poder interactuar y lo ejemplifican jugando con la iluminación o incluso inundando la escena.

Una de las limitaciones que tiene este método es que está diseñado para fotografía en 360° y no se menciona en ningún momento al vídeo 360 por lo que se puede deducir que no está preparado. Por otro lado este tipo de procesamiento de imágenes requiere una cantidad grande de tiempo por lo que es posible que no esté al alcance de todos.

[3]

6. Desarrollo

aasd asd asdafew ferg dfg sg

7. Resultados

aasd asd asdafew ferg dfg sg

8. Conclusiones

aasd asd asdafew ferg dfg sg

Bibliografía

- [1] Adobe Creative Cloud. projectsidewinder: Adobe max 2017 (sneak peeks) | adobe creative cloud.
- [2] Josh Gladstone. Use depth maps to create 6 dof in unity.
- [3] Jingwei Huang, Zhili Chen, Duygu Ceylan, and Hailin Jin. 6-dof vr videos with a single 360-camera. *2017 IEEE Virtual Reality (VR)*, pages 37–44, 2017.
- [4] Facebook Inc. Casual 3d capture.
- [5] UploadVR. Point cloud vr - htc vive.