



Facultad de Informática de la Universidad de A Coruña
Electrónica y Sistemas

PROYECTO DE FIN DE CARRERA
INGENIERIA TÉCNICA EN INFORMÁTICA DE SISTEMAS

Visualización avanzada de nubes de puntos con OpenGL

Alumno: Antúnez González, David
Director: Álvarez Mures, Luis Omar
Tutor y Director: Padrón González, Emilio José
Fecha: 5 de septiembre de 2015

D. NOMBRE DEL DIRECTOR DEL PROYECTO
Profesor de Escuela o Facultad Universitaria
Departamento de lo que sea
Universidad de A Coruña

CERTIFICA: Que la memoria titulada “*Lorem ipsum dolor sit amet, consectetuer adipiscing elit vestibulum a pede sit amet augue pellentesque*” ha sido realizada por MAURO SILVOSA RIVERA bajo mi dirección y constituye su Proyecto de Fin de Carrera de Nombre de la Titulación Que Corresponda.

En A Coruña, a 5 de septiembre de 2015

D. NOMBRE DEL DIRECTOR DEL PROYECTO
Director del proyecto

Resumen:

El render de nubes de puntos adquirió un renovado interés en los últimos años con la popularización y proliferación de nuevos sistemas de adquisición de datos de medio coste tipo LiDAR o de bajo basados en fotogrametría o cámaras infrarrojas.

Estos dispositivos obtienen una nube de puntos 3D (posición geométrica), con información adicional arbitraria asociada, normalmente el color como mínimo, pero también es posible otros datos como normales, temperatura, etc. Por las propias características que tienen el punto como primitiva gráfica, respecto por ejemplo de los tradicionales polígonos (no tienen área, no tienen orientación, no están conectados, etc.), el render o visualización de nubes de puntos presenta una serie de retos si queremos obtener una visualización realista y de calidad.

Este proyecto se centra en la visualización avanzada de nubes de puntos, aplicando algunas de las más novedosas técnicas de render y haciendo uso de las características avanzadas de la API gráfica multiplataforma OpenGL y explotando el hardware de las tarjetas gráficas modernas. El resultado del proyecto será la implementación de una herramienta multiplataforma para la visualización avanzada de nubes de puntos 3D.

Lista de palabras clave:

 Lorem ipsum, dolor sit amet, consectetuer, adipiscing elit, vestibulum a pede, sit amet augue, pellentesque ,vehicula, donec pretium.

Knight Rider, a shadowy flight into the dangerous world of a man who does not exist. Michael Knight, a young loner on a crusade to champion the cause of the innocent, the helpless in a world of criminals who operate above the law.

Agradecimientos¹

A D./Dña. Nombre De La Persona (Departamento u organización, si corresponde), por descripción de la ayuda prestada, si interesa resaltarlo.

A D./Dña. Nombre De La Persona (Departamento u organización, si corresponde), por descripción de la ayuda prestada, si interesa resaltarlo.

A D./Dña. Nombre De La Persona (Departamento u organización, si corresponde), por descripción de la ayuda prestada, si interesa resaltarlo.

¹En orden alfabético

Índice general

Introducción	xI
1. Planificación y metodología	1
1.1. Desarrollo ágil de software	1
1.1.1. Agile Manifesto	2
1.1.2. Principios Ágiles	2
1.1.3. Aplicación	3
2. Fundamentos tecnológicos	8
2.1. Nubes de puntos	8
2.1.1. Adquisición	9
2.2. OpenGL	10
2.2.1. Pipeline	10
2.2.2. Shaders	12
2.3. PCL	13
3. Fundamentos de la Programación Gráfica	15
3.1. Render	15
3.1.1. Modelos 3D	17
3.2. Espacios 3D Euclídeos y transformaciones	17
3.3. Camera model	19
3.3.1. Parámetros de la Cámara	20

3.3.2. Funcionamiento interno de la cámara	21
4. Visualización de nubes de puntos con GPU	25
4.1. Técnicas de rasterización	25
4.1.1. Sized-Fixed Points	26
4.1.2. Image-aligned Squares	26
4.1.3. Affinely Projected Point Sprites	27
4.1.4. Perspective Correct Rasterization	28
4.2. Técnicas de Blending	30
4.2.1. Flat Shading	30
4.2.2. Gouraud Shading	32
4.2.3. Phong Shading	34
4.2.4. Deferred Shading	35
5. Resultados y rendimiento	37
6. Conclusiones	38

Índice de figuras

2.1.	Mapa de desviaciones de una pieza con respecto a un modelo <i>CAD</i> en un proceso industrial.	9
2.2.	Estación Leica LIDAR utilizado para el escaneo de edificios, formaciones rocosas, etc. con el objetivo de generar modelos 3D.	11
2.3.	El pipeline de OpenGL, las fases es negrita son programables.	12
2.4.	Evolución del número de transistores en GPU y CPU (1999 - 2010).	13
3.1.	Fotorealistic (izquierda) vs. No fotorealísticos (derecha).	16
3.2.	Funcionamiento de una cámara estenopeica.	19
3.3.	Representación de las diferencias entre cámara en perspectiva (izquierda) y ortográfica (derecha).	21
3.4.	La matriz P se encarga de transformar la pirámide de visualización en un cubo unitario.	24
4.1.	Todos los puntos de la nube son renderizados con un cuadrado del mismo tamaño para toda la nube.	26
4.2.	Dragon, modelo de 845,281 puntos renderizado mediante Affinely Projected Point Sprites. Este método causa huecos en ángulos extremos.	29
4.3.	Proyección y Ray Casting.	30
4.4.	Suzanne, modelo de 7,958 puntos renderizado mediante Perspective Correct Rasterization.	31
4.5.	Lucy, modelo de 397,664 puntos. Renderizado mediante <i>Flat Shading</i> (izquierda), <i>Gouraud Shading</i> (centro), y <i>Phong Shading</i> (derecha).	31
4.6.	Los tres pasos, <i>visibility</i> , <i>blending</i> , y <i>normalization</i>	33

4.7. Los vectores n_i son representados como puntos homogéneos $(x,y,1)$ en un plano desplazado tangente.	34
4.8. El renderizado mediante <i>Deferred Shading</i> acumula atributos como color y normales, seguido por una normalización y un shading pass.	35

Índice de cuadros

1.1. Estimación de coste	7
------------------------------------	---

Introducción

One for all and all for one, Muskehounds are always ready. One for all and all for one, helping everybody. One for all and all for one, it's a pretty story. Sharing everything with fun, that's the way to be. One for all and all for one, Muskehounds are always ready. One for all and all for one, helping everybody. One for all and all for one, can sound pretty corny. If you've got a problem chum, think how it could be.

Hong Kong Phooey, number one super guy. Hong Kong Phooey, quicker than the human eye. He's got style, a groovy style, and a car that just won't stop. When the going gets tough, he's really rough, with a Hong Kong Phooey chop (Hi-Ya!). Hong Kong Phooey, number one super guy. Hong Kong Phooey, quicker than the human eye. Hong Kong Phooey, he's fan-riffic!

Ten years ago a crack commando unit was sent to prison by a military court for a crime they didn't commit. These men promptly escaped from a maximum security stockade to the Los Angeles underground. Today, still wanted by the government, they survive as soldiers of fortune. If you have a problem and no one else can help, and if you can find them, maybe you can hire the A-team.

This is my boss, Jonathan Hart, a self-made millionaire, he's quite a guy. This is Mrs H., she's gorgeous, she's one lady who knows how to take care of herself. By the way, my name is Max. I take care of both of them, which ain't easy, 'cause when they met it was MURDER!

Top Cat! The most effectual Top Cat! Who's intellectual close friends get to call him T.C., providing it's with dignity. Top Cat! The indisputable leader of the gang. He's the boss, he's a pip, he's the championship. He's the most tip top, Top Cat.

Contexto y Objetivos

Los objetivos concretos que cubre la elaboración de este proyecto fin de carrera son:

1. Aprendizaje de la API gráfica OpenGL, que permite un render interactivo multiplataforma.
2. Visualización básica: Desarrollo de un visualizador multiplataforma de nubes de puntos 3D. Se hará una primera versión básica que irá evolucionando durante todo el proyecto, incorporando las distintas técnicas que se vayan estudiando e implementando.
3. Implementación de técnicas avanzadas de visualización de nubes de puntos
 - a) Basadas en rasterización: discos orientados, corrección de perspectiva
 - b) Basadas en blending
4. Propuesta de un algoritmo novedoso para la visualización de puntos con blending en un único pase.
5. Integración del visualizador con la libtería PCL, especializada en la gestión de nubes de puntos. Esto nos permitirá acceder a un importante abanico de métodos y estructuras para el trabajo eficiente con grandes conjuntos de datos.

Estructura de la memoria

In a dolor sed odio eleifend varius. Nam ullamcorper. Curabitur ut erat vulputate nisi molestie tempus. Sed aliquam rutrum odio. In mollis. Fusce consectetuer lorem nec diam. Sed mollis lacinia purus. Curabitur feugiat hendrerit neque. Quisque auctor laoreet diam. Curabitur sit amet nisi. Fusce velit massa, dignissim quis, bibendum eget, vehicula mattis, leo. Morbi auctor leo sit amet nibh. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Nullam enim. Pellentesque hendrerit, augue non vulputate semper, sem lorem pharetra nibh, sit amet egestas massa diam ac augue. In dui nulla, egestas nec, pulvinar suscipit, tincidunt ornare, nisi. Duis tristique tortor quis magna. Vestibulum faucibus lorem nec neque. Sed nec nibh. Nunc condimentum. Maecenas neque. Nullam pretium est non risus. Etiam gravida. Maecenas nisl. Fusce pharetra odio in tortor. Integer orci turpis, interdum eget, vulputate sed, tristique a, metus. Duis vitae dui quis lectus pretium aliquam. Praesent quam.

Aliquam sed orci. Cras adipiscing nisl quis pede. Ut rhoncus. Donec viverra laoreet purus. Phasellus nulla. Vivamus eget eros. In mollis aliquam orci. Proin ullamcorper. Nullam sollicitudin vestibulum lorem. Nunc malesuada sagittis augue. Donec tellus velit, dapibus a, aliquam ac, tincidunt id, lectus.

Capítulo 1. Phasellus tempor velit nec velit. Proin vitae dui a sapien commodo blandit. Etiam aliquam, sapien vitae fringilla venenatis, lectus sem accumsan orci, eget blandit orci odio et magna. Quisque malesuada, eros vel tempus eleifend, velit enim porttitor sem, eget consequat nulla neque et sapien. Morbi leo. Sed vestibulum lacus. Fusce ut lacus. Phasellus pellentesque pede eu eros. Duis turpis felis, eleifend ut, semper ac, porta nec, sem. Praesent odio. Sed laoreet mollis purus. Praesent vestibulum, velit ut mollis aliquam, quam lectus varius urna, sed ultricies erat nisl ac tortor. Vivamus tempor mauris sit amet nulla. Integer venenatis. Integer sagittis euismod ante. Suspendisse at elit. Duis eget purus nec pede adipiscing auctor. Proin ac est.

Capítulo 2. Proin condimentum. Maecenas sodales. In ornare nunc a leo. Nam sit amet ligula. Nunc quis urna ac metus imperdiet lobortis. Sed quis ligula. Maecenas blandit pede. Donec lacinia rutrum ligula. Vivamus in metus vel elit pharetra molestie.

Capítulo 3. Nullam ante lorem, placerat et, egestas nec, pellentesque non, sapien. Donec semper, felis id posuere faucibus, nibh ipsum tincidunt quam, et varius ipsum odio ac neque. In tincidunt dignissim diam. Sed lacus lorem, ornare ut, eleifend vel, pellentesque tempus, augue. Duis eu magna. Mauris libero ante, porttitor vel, lobortis a, mollis ac, sem. Nunc at lectus. Integer ac libero a nisl dignissim mollis. Donec velit neque, vestibulum eget, pulvinar vel, malesuada ut, nisi. Praesent congue tempus quam. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus.

Capítulo 4. Mauris ut odio. Nulla accumsan. Morbi condimentum fermentum purus. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Nunc dignissim, neque eget convallis pretium, diam tortor fringilla lacus, a laoreet nisl metus eu magna. Cras ut lectus. Etiam accumsan feugiat elit.

Capítulo ... Donec a pede. Proin dolor. Ut nunc ligula, tempor id, ornare sit amet, aliquam et, nibh. In mollis iaculis pede. Vivamus gravida orci eu nisl. Sed nibh sem, consequat at, iaculis non, placerat in, ligula. Praesent id nisi. Nunc pellentesque justo non libero. Sed quis est sit amet purus lobortis blandit. Sed arcu justo, rhoncus condimentum, ullamcorper iaculis, viverra et, nisl.

Capítulo N. Fusce luctus gravida leo. Nullam dignissim arcu ac risus hendrerit rhoncus. Aliquam erat volutpat. Ut mollis, mauris non aliquam luctus, nulla sem aliquam tellus, in consequat augue odio in urna.

Capítulo 1

Planificación y metodología

Este capítulo explica el método escogido para el desarrollo de este proyecto. Ya que fundamentalmente tiene una base de carácter investigativa, se pensó que encajaría perfectamente un sistema ágil como el que a continuación se comentará para el desarrollo del proyecto. Seguidamente de la exposición de un diagrama de *Gantt* con los plazos e hitos que se siguieron en este caso concreto.

1.1. Desarrollo ágil de software

El desarrollo ágil de software refiere a métodos de ingeniería del software basados en el desarrollo iterativo e incremental, donde los requisitos y soluciones evolucionan mediante la colaboración de grupos auto organizados y multidisciplinarios. Existen muchos métodos de desarrollo ágil; la mayoría minimiza riesgos desarrollando software en lapsos cortos. El software desarrollado en una unidad de tiempo es llamado una iteración, la cual debe durar de una a cuatro semanas. Cada iteración del ciclo de vida incluye: planificación, análisis de requisitos, diseño, codificación, revisión y documentación. Una iteración no debe agregar demasiada funcionalidad para justificar el lanzamiento del producto al mercado, sino que la meta es tener una «demo» (sin errores) al final de cada iteración. Al final de cada iteración el equipo vuelve a evaluar las prioridades del proyecto.

Los métodos ágiles enfatizan las comunicaciones cara a cara en vez de la documentación. La mayoría de los equipos ágiles están localizados en una simple oficina abierta, a veces llamadas "plataformas de lanzamiento" (bullpen en inglés). La oficina debe incluir revisores, escritores de documentación y ayuda, diseñadores de iteración y directores de proyecto. Los métodos ágiles también enfatizan que el software funcional es la primera medida del progreso. Combinado con la preferencia por las comunicaciones cara a cara,

generalmente los métodos ágiles son criticados y tratados como "indisciplinados" por la falta de documentación técnica.

Los métodos de desarrollo ágiles e iterativos pueden ser vistos como un retroceso a las prácticas observadas en los primeros años del desarrollo de software (aunque en ese tiempo no había metodologías formales). Inicialmente, los métodos ágiles fueron llamados métodos de "peso liviano".

1.1.1. Agile Manifesto

En el año 2001, miembros prominentes de la comunidad se reunieron en Snowbird, Utah, para discutir métodos de desarrollo ligero, publicando lo que denominaron *Manifesto for Agile Software Development* con motivo de sentar las bases de lo que ahora se conoce como Desarrollo ágil de software.

Siendo sus ideas principales:

- **Individuos e interacciones** auto-organización y la motivación son importantes. Otros valores promovidos son la co-ubicación y la programación en parejas.
- **Software funcionando** priorizando el tener *demos* funcionales frente a documentación.
- **Colaboración con el cliente** los requerimientos técnicos muchas veces no pueden ser definidos al principio del ciclo de desarrollo, por lo que la continua participación con el cliente es muy importante
- **Respuesta ante el cambio** centrándose en respuestas rápidas a los cambios y el desarrollo continuo.

1.1.2. Principios Ágiles

Este manifesto está basado en los siguientes 12 principios:

- Nuestra mayor prioridad es satisfacer al cliente mediante la entrega temprana y continua de software con valor.
- Aceptamos que los requisitos cambian, incluso en etapas tardías del desarrollo. Los procesos Ágiles aprovechan el cambio para proporcionar ventaja competitiva al cliente.

- Entregamos software funcional frecuentemente, entre dos semanas y dos meses, con preferencia al periodo de tiempo más corto posible.
- Los responsables de negocio y los desarrolladores trabajamos juntos de forma cotidiana durante todo el proyecto.
- Los proyectos se desarrollan en torno a individuos motivados. Hay que darles el entorno y el apoyo que necesitan, y confiarles la ejecución del trabajo.
- El método más eficiente y efectivo de comunicar información al equipo de desarrollo y entre sus miembros es la conversación cara a cara.
- El software funcionando es la medida principal de progreso.
- Los procesos Ágiles promueven el desarrollo sostenible. Los promotores, desarrolladores y usuarios debemos ser capaces de mantener un ritmo constante de forma indefinida.
- La atención continua a la excelencia técnica y al buen diseño mejora la Agilidad.
- La simplicidad, o el arte de maximizar la cantidad de trabajo no realizado, es esencial.
- Las mejores arquitecturas, requisitos y diseños emergen de equipos auto-organizados.
- A intervalos regulares el equipo reflexiona sobre cómo ser más efectivo para a continuación ajustar y perfeccionar su comportamiento en consecuencia.

1.1.3. Aplicación

En gran parte del proyecto, fueron aplicadas premisas e ideas propuestas en este *Agile Manifesto*. Se hicieron reuniones semanales en las que se discutían las siguientes tareas a llevar a cabo, dividiendo objetivos principales en tareas simples que no supusiesen más de 2 semanas de desarrollo. El trabajo en cada iteración era constante, centrándose siempre en completar la tarea y no asumiendo cambios si no eran profundamente necesario.

El diseño incremental fué presente, puesto que en cada momento el diseño era revisado para adaptarlo a los nuevos requerimientos.

Las revisiones con mis dos directores fueron siempre cara a cara mientras fue posible. En cada reunión se comentaba que objetivos habían sido cumplidos desde la última vez, y que objetivos se intentarían abordar en la siguiente iteración. Cuando estas reuniones no eran posibles, sistemas como el correo electrónico o plataformas de *chat* como *Hangouts* o *Telegram* fueron utilizados.

Para mantener el código accesible en todo el momento, desde el comienzo del desarrollo se utilizaron sistemas de control de versiones como *Git*. Que es un sistema libre de administración de código fuente y distribución. De modo que además de servir como un sistema de *backup* era fácil llevar el seguimiento del desarrollo, facilitando la creación de nuevas ramas paralelas a la versión estable, con motivo de prueba para nuevos algoritmos.

A pesar de que la naturaleza investigativa de este proyecto imposibilita la buena estimación de tiempo por cada iteración, el siguiente diagrama de *Gantt* viene a resumir los tiempos que llevó cada iteración.

Tarea	Fecha de inicio	Fecha de fin
Creación de prototipo	1/05/15	5/05/15
Configuración del entorno de trabajo + CMake	1/05/15	1/05/15
Inicialización de contexto	4/05/15	4/05/15
Test de OpenGL + Shader básico	5/05/15	5/05/15
Inclusión de una cámara orbital	6/05/15	7/05/15
Instalación de GLM	6/05/15	6/05/15
Implementación de la cámara	7/05/15	7/05/15
Sized-Fixed Point	8/05/15	11/05/15
Generación de buffers en la GPU	8/05/15	8/05/15
Clase VAO + Shader Sized-Fixed Point	11/05/15	11/05/15
Port de GLUT a GLFW	12/05/15	12/05/15
Image-aligned Squares Shader	13/05/15	19/05/15
Sistema de cambio de Shader	20/05/15	20/05/15
Affinely Projected Point Sprites Shader	21/05/15	25/05/15
Soporte con PCL	26/05/15	27/05/15
Carga de formatos .PLY y .PCD	26/05/15	26/05/15
Refactorización hacia estructura cloud de PCL	26/05/15	27/05/15
Perspectively Correct Rasterization Shader	28/05/15	10/06/15
Radio del splat variable según vecindad	11/06/15	12/06/15
Gouraud Shading	15/06/15	2/07/15
Añadir al pipeline opción a sistema MultiPass	15/06/15	18/06/15
Gouraud Shading Shader	19/06/15	2/07/15
Implementación de FXAA	15/06/15	2/07/15
Gouraud Shading	3/07/15	6/07/15
Phong Shading Shader	7/07/15	13/07/15
Deferred Shading Shader	14/07/15	22/07/15
Luz orbital	23/07/15	23/07/15

1.1. Desarrollo ágil de software

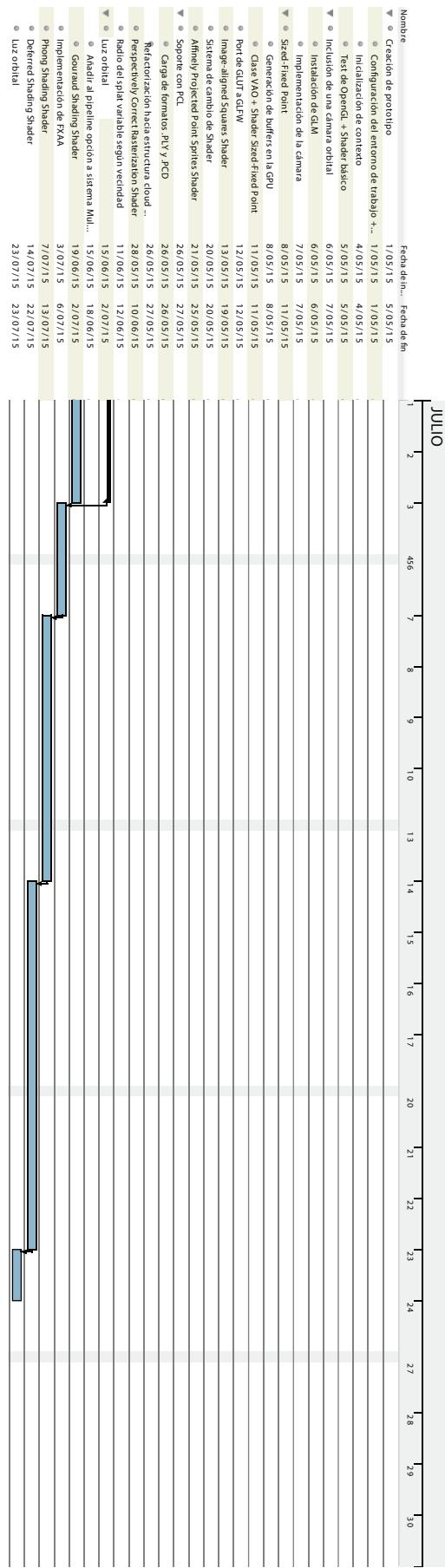
MAYO

Nombre	Fecha de in...	Fecha de fin
● Creación de prototipo	1/05/15	5/05/15
○ Configuración del entorno de trabajo +..	1/05/15	1/05/15
○ Inicialización de contexto	4/05/15	4/05/15
○ Test de OpenGL + Shader básico	5/05/15	5/05/15
○ Inclusión de una cámara orbital	6/05/15	7/05/15
○ Instalación de GLM	6/05/15	6/05/15
○ Implementación de la cámara	7/05/15	7/05/15
○ Sized-Fixed Point	8/05/15	11/05/15
○ Generación de buffer en la GPU	8/05/15	8/05/15
○ Clase AOO + Shader Sized-Fixed Point	11/05/15	11/05/15
○ Port de GLUT y GLFW	12/05/15	12/05/15
○ Image-Aligned Square Shader	13/05/15	19/05/15
○ Sistema de cambio de Shader	20/05/15	20/05/15
○ Affinity Projected Dumb Sprites Shader	21/05/15	22/05/15
○ Soporte con PCL	26/05/15	27/05/15
○ Carga de formatos PLY y PCD	26/05/15	26/05/15
○ Refactorización hacia estructura cloud...	26/05/15	27/05/15
○ PerspectiveCorrect Rasterization Shader	28/05/15	10/06/15
○ Ratio del spilt variable según vecindad	11/06/15	12/06/15
○ Gouraud Shading	15/06/15	2/07/15
○ Añadir al pipeline opción a sistemas Mul... -	18/06/15	18/06/15
○ Gouraud Shading Shader	19/06/15	2/07/15
○ Implementación de FXA	3/07/15	6/07/15
○ Phong Shading Shader	7/07/15	13/07/15
○ Deferred Shading Shader	14/07/15	22/07/15
○ Luz orbital	23/07/15	27/07/15

JUNIO

Nombre	Fecha de in...	Fecha de fin
▼ ○ Creación de prototipo	1/05/15	5/05/15
○ Configuración del entorno de trabajo +...	1/05/15	1/05/15
○ Inicialización de contexto	4/05/15	4/05/15
○ Test de OpenGL + Shader básico	5/05/15	5/05/15
▼ ○ Inclusión de una cámara orbital	6/05/15	7/05/15
○ Instalación de GLM	6/05/15	6/05/15
○ Implementación de la cámara	7/05/15	7/05/15
▼ ○ Sized-Fixed Point	8/05/15	11/05/15
○ Generación de buffer en la GPU	8/05/15	8/05/15
○ Clase AOO + Shader Sized-Fixed Point	11/05/15	11/05/15
○ Port de GLUT y GLFW	12/05/15	12/05/15
○ Image-Aligned Square Shader	13/05/15	19/05/15
○ Sistema de cambio de Shader	20/05/15	20/05/15
○ Affinity Projected Dumb Sprites Shader	21/05/15	22/05/15
○ Soporte con PCL	26/05/15	27/05/15
○ Carga de formatos PLY y PCD	26/05/15	26/05/15
○ Refactorización hacia estructura cloud...	26/05/15	27/05/15
○ PerspectiveCorrect Rasterization Shader	28/05/15	10/06/15
○ Ratio del spilt variable según vecindad	11/06/15	12/06/15
○ Gouraud Shading	15/06/15	2/07/15
○ Añadir al pipeline opción a sistemas Mul... -	18/06/15	18/06/15
○ Gouraud Shading Shader	19/06/15	2/07/15
○ Implementación de FXA	3/07/15	6/07/15
○ Phong Shading Shader	7/07/15	13/07/15
○ Deferred Shading Shader	14/07/15	22/07/15
○ Luz orbital	23/07/15	27/07/15

1.1. Desarrollo ágil de software



Recurso	Cantidad	Horas	Coste por hora	Total
Software				0€
Ordenador	1	480	0,1372€	65,86€
Analista-Programador	1	480	15€	7200€
				7265,86€

Cuadro 1.1: Estimación de coste

Se tiene además un presupuesto para un proyecto como este. Incluyendo únicamente costes humanos y hardware, ya que el *software* fué siempre gratuito. El precio por hora del ordenador fue calculado pensando en una amortización de 5 años fiscales para el equipo.

Capítulo 2

Fundamentos tecnológicos

En este capítulo se explicarán los fundamentos tecnológicos necesarios para la comprensión de este proyecto. Empezando por la explicación de las Nubes de puntos (ver 2.1), OpenGL (ver 2.2) y la librería PCL (ver 2.3).

2.1. Nubes de puntos

Una nube de puntos es un conjunto de puntos en un sistema de coordenadas tridimensional. Estos puntos se identifican habitualmente como coordenadas X, Y, y Z y son representaciones de la superficie externa de un objeto.

Las nubes de puntos tienen múltiples aplicaciones, entre las que se incluyen la elaboración de modelos tridimensionales en CAD de piezas fabricadas, la inspección de calidad en metrología, y muchas otras en el ámbito de la visualización, animación, texturización y aplicaciones de personalización masiva.

Aunque las nubes de puntos se pueden revisar y texturizar directamente, habitualmente no se utilizan de esta forma en la mayoría de las aplicaciones tridimensionales, ya que se convierten en modelos de mallas poligonales o mallas triangulares irregulares, modelos de superficie *NURBS*, o modelos de *CAD* mediante un proceso denominado reconstrucción de superficies.

Una aplicación en la que las nubes de puntos se utilizan directamente es la metrología y en la inspección industrial (ver Figura 2.1). La nube de una pieza fabricada se alinea a un modelo CAD (o a otra nube de puntos) y se comparan para verificar las diferencias. Éstas se visualizan como mapas de color que proporcionan un indicador visual de la desviación entre la pieza fabricada y el modelo CAD. Las dimensiones geométricas y tolerancias

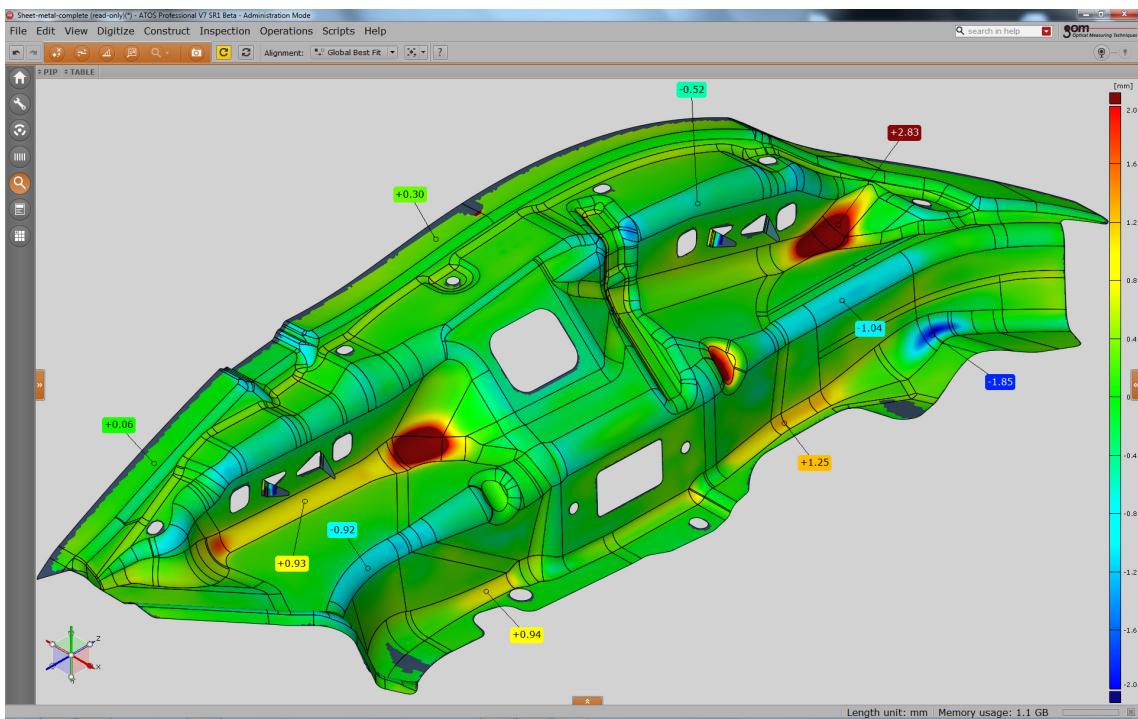


Figura 2.1: Mapa de desviaciones de una pieza con respecto a un modelo *CAD* en un proceso industrial.

también se pueden extraer directamente de la nube de puntos.

Las nubes de puntos también se pueden utilizar para representar datos volumétricos como en la imagen médica. Se logra usando nubes de puntos de muestreo múltiple y compresión de datos.

En los Sistemas de Información Geográfica *GIS* las nubes de puntos son una de las fuentes de datos para construir modelos digitales del terreno.

2.1.1. Adquisición

Para la adquisición de estas nubes de puntos, existen actualmente diferentes métodos:

- **Escaner 3D** Un escáner 3D es un dispositivo que analiza un objeto o una escena para reunir datos de su forma y ocasionalmente su color. Diferenciando dos tipos de escáneres en función de si existe contacto con el objeto o no y dividiendo estos últimos en otras dos categorías (activos y pasivos).
- **Contacto** Los escáneres 3D examinan el objeto apoyando el elemento de medida (palpador) sobre la superficie del mismo, típicamente una punta de acero

duro o zafiro. Una serie de sensores internos permiten determinar la posición espacial del palpador. Se usan en su mayoría en control dimensional en procesos de fabricación y pueden conseguir precisiones típicas de 0,01 mm.

- **Sin contacto**

- **Activos** Los escáneres activos (ver Figura 2.2) emiten alguna clase de señal y analizan su retorno para capturar la geometría de un objeto o una escena. Se utilizan radiaciones electromagnéticas (desde ondas de radio hasta rayos X) o ultrasonidos.
- **Pasivos** Los escáneres pasivos no emiten ninguna clase de radiación por sí mismos, pero en lugar se fía de detectar la radiación reflejada del ambiente. La mayoría de los escáneres de este tipo detectan la luz visible porque es una radiación ya disponible en el ambiente. Otros tipos de radiación, tal como el infrarrojo podrían ser utilizados también. Los métodos pasivos pueden ser muy baratos, porque en la mayoría de los casos estos no necesitan hardware particular.

- **Fotogrametría** La fotogrametría es una técnica para determinar las propiedades geométricas de los objetos y las situaciones espaciales a partir de imágenes fotográficas. Básicamente, es una técnica de medición de coordenadas 3D, que utiliza fotografías u otros sistemas de percepción remota junto con puntos de referencia, como medio fundamental para la medición.

2.2. OpenGL

OpenGL (Open Graphics Library) es una librería multiplataforma destinada a renderizado 2D y 3D. Esta API es típicamente usada para trabajar con la Unidad Gráfica de Procesado (GPU) para conseguir un renderizado acelerado mediante *hardware*, aunque es posible que esta sea implementada enteramente mediante *software*.

2.2.1. Pipeline

El *Pipeline* de render es una secuencia de pasos funcionales para la obtención de imágenes de escenas 3D. Idealmente, si se divide un proceso en n etapas se incrementará la velocidad del proceso en ese factor n . Así, la velocidad de la cadena viene determinada por el tiempo requerido por la etapa más lenta.



Figura 2.2: Estación Leica LIDAR utilizado para el escaneo de edificios, formaciones rocosas, etc. con el objetivo de generar modelos 3D.

El *pipeline* que OpenGL sigue en la renderización de objetos (ver Figura 2.3) sería el siguiente:

1. Preparación del *Vertex Array Object* a renderizar.
2. Procesamiento de los vértices:
 - a) Cada vértice del flujo de datos (*stream*) es procesado mediante el *vertex shader* con motivo de realizar las transformaciones necesarias de estos a un espacio de coordenadas adaptado a su proyección.
 - b) Fase opcional de teselación.
 - c) Fase opcional *geometry shader*. La salida es una secuencia de primitivas.
3. *Vertex Post-Processing*, la salida del anterior paso es ajustado (*primitive clipping* y transformación del *viewport* a coordenadas de la ventana).
4. *Primitive Assembly*.
5. Las primitivas son divididas en elementos discretos, los cuales generan un número determinado de fragmentos.

6. El *fragment shader* procesa cada fragmento generando un numero diferente de salidas.
7. *Per-Sample Processing*, donde se descartan fragmentos, se combinan los colores de los fragmentos obtenidos del *fragment shader* con los que ya existen en el *buffer*, ademas de realizar diferentes operaciones lógicas.

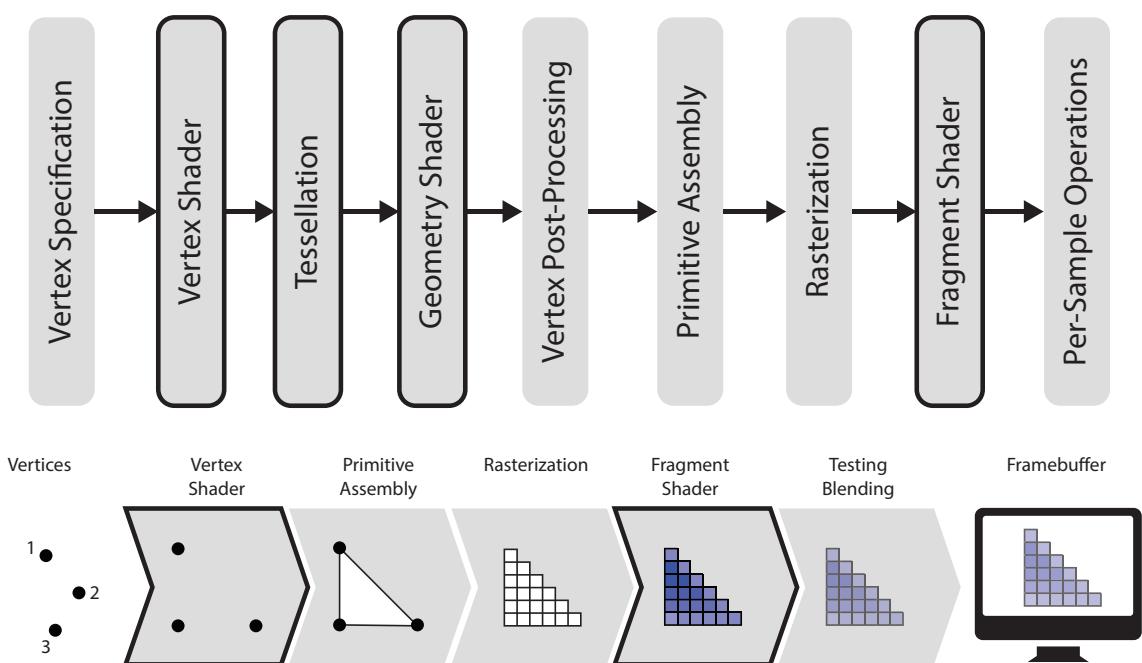


Figura 2.3: El pipeline de OpenGL, las fases es negrita son programables.

2.2.2. Shaders

El *hardware* de aceleración gráfica ha sufrido una importante transformación en la última década. Como se muestra en la Figura 2.4, en los últimos años el potencial de las *GPUs* ha superado con creces al de la *CPU*.

Esta tendencia en el *hardware* gráfico ha reemplazado la funcionalidad fija del antiguo *pipeline* con areas programables mediante pequeños programas. De este modo algunas de las fases del *pipeline* pueden ser modificadas con pequeños programas escritos por el usuario denominados *shaders*. Estos programas se ejecutan directamente en la *GPU*, y permite realizar operaciones a diferentes niveles con alta eficiencia.

El lenguaje de OpenGL de *Shading (GLSL)* está basado en ANSI C y muchas de las características han sido mantenidas excepto las que entraban en conflicto con temas de

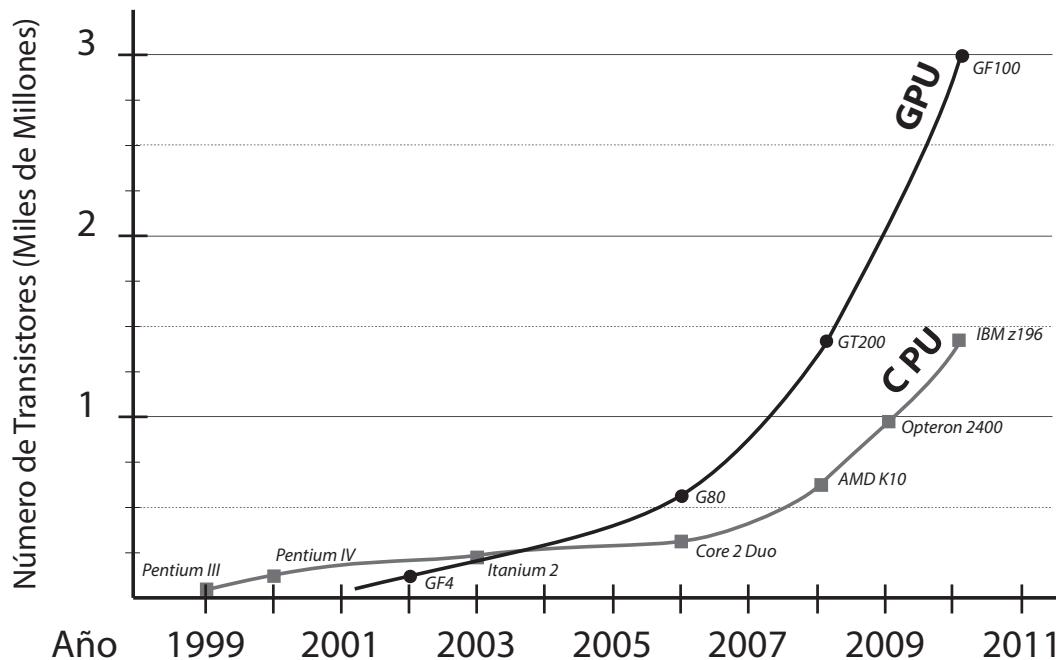


Figura 2.4: Evolución del número de transistores en GPU y CPU (1999 - 2010).

rendimiento o implementación. C fue extendido con tipos nuevos para manejar estructuras como vectores o matrices, para hacer más fácil operaciones típicas en gráficos 3D.

En la figura 2.3 se puede observar que en OpenGL las fases referentes al *vertex shader*, *tessellation*, *geometry shader* o *fragment shader* pueden ser programables.

2.3. PCL

Point Cloud Library (PCL) es una librería *open-source* con algoritmos destinados a tareas de procesamiento de nubes de puntos y geometría 3D. La librería contiene algoritmos para estimación, reconstrucción de superficies, *model fitting* y segmentación. Está escrita en C++ y distribuida bajo una licencia *BSD*.

Estos son usados, por ejemplo, para percepción en robótica para filtrar ruido, unir nubes de puntos 3D, extraer *keypoints* y calcular descriptores para reconocimiento de objetos basados en su geometría y crear superficies a partir de nubes de puntos para visualizarlas.

El desarrollo de esta librería comenzó en Marzo de 2010 The development of the Point Cloud Library started in March 2010 en la empresa *Willow Garage*. La primera versión

oficial de *PCL* fue lanzada en Mayo de 2011.

Capítulo 3

Fundamentos de la Programación Gráfica

En este capítulo se introducirán algunos conceptos involucrados en el *render* 3D.

3.1. Render

Tradicionalmente la Programación Gráfica ha sido definida como la disciplina dedicada a sintetizar imágenes algorítmicamente mediante ordenadores. Actualmente se pueden encontrar más temas acerca del hiperrealismo, técnicas de animación, realidad virtual, etc. Para generar imágenes a partir de escenar tridimensionales un proceso denominado *render* es usado. Este conjunto de acciones está al cargo del modelado y sus propiedades, iluminación y si es necesario la cámara que capturará todo.

Como bien se dijo antes, el *render* es el proceso computacional de generar una imagen a partir de un modelo. Desde esta definición, se podría sacar diferentes interpretaciones, desde crear una película de animación 3D a un diagrama de barras, y todas ellas serían igualmente válidas. Aunque el término que normalmente es usado es referido a un modelo tiene una naturaleza espacial y más específicamente tridimensional.

Diferentes clasificaciones de técnicas de *render* puede ser definidas, pero para esta introducción usaremos dos de ellas. Por un lado, si nos fijamos en el estilo de la imagen que se pretende conseguir:

- **Fotorealístico** que intenta ser lo más fiel como sea posible a la realidad. Este tipo de *render* puede ser subdividido en:



Figura 3.1: Fotorealistic (izquierda) vs. No fotorealísticos (derecha).

- **Render basado en Físicas** Intentando calcular una simulación lo más autética posible del tratamiento de la luz a través de la escena y su interacción con materiales y objetos. La precisión de la simulación dependerá de los modelos físicos y matemáticos escogidos.
- **Falseados** que utilizan trucos para reducir tiempos de renderizado.
- **No fotorealísticos** que pretende usar otros tipos de efectos para la renderización de la escena con motivos artísticos.

La Figura 3.1 muestra la diferencia entre ambas aproximaciones.

Por otro lado, si se comparan las opciones de interacción de la aplicación con el usuario final, se podría dividir las técnicas en 2 apartados:

- **Render Offline** Donde el proceso de generar la imagen es demasiado lento para responder instantáneamente a las interacciones del usuario. Una escala de tiempo de segundos a incluso días podría considerarse "lento". Esto ocurre por ejemplo cuando se generan fotogramas para una película.
- **Render Online o Reak-time** Donde el tiempo para la obtención de las imágenes es suficientemente corto como para responder a la interacción del usuario en la aplicación, teniendo este una sensación de continuidad. Una escala de *milisegundos* tendría que ser alcanzada para obtener este efecto. Normalmente medida en *frames per second* (FPS). Un buen ejemplo de este grupo podrían ser los videojuegos.

La mayor parte del trabajo realizado en este documento es centrado en conseguir un render online, incluso si esto significa sacrificar un mejor aproximación fotorealística.

3.1.1. Modelos 3D

El modelado describe el proceso de formar la figura de un objeto virtual en un ordenador. Existen realmente tres tipos diferentes de Modelos 3D:

- **Basado en Polígonos:** Donde se describe la superficie de un objeto mediante un conjunto de polígonos. La forma triangular es la mas comunmente usada para esta tarea, puesto que son planos y trivialmente convexos. Practicamente la totalidad del *hardware* gráfico esta optimizado pensando en esta primitiva.
Aunque en el proceso del modelado de un objeto, este no tiene porque estar construido usando triangulos. En última instancia, es convertido a un conjunto de triángulos mediante un proceso conocido como teselación.
- **Basado en Voxels:** Divide el espacio tridimensional en una rejilla donde el *voxel* es la mínima unidad de medida. Cada una de estas celdas es rellenada o no, dependiendo de si se pretende que sea renderizada. La memoria necesaria aumenta en función de la precisión necesaria. Este tipo de procesos usualmente utilizado en Informática Biomédica.
- **Basada en Puntos:** Los objetos son representados mediante puntos que son producto de un muestreo de su superficie. Cada punto tiene una posición y alguna información relativa a la superficie a la que pertenece. Comparado con el modelo tradicional basado en polígonos, esta primitiva necesita sus propias técnicas de *render*. Este es nuestro caso de estudio.

3.2. Espacios 3D Euclídeos y transformaciones

En gráficos por computación normalmente trabajamos con espacios tridimensionales de geometría Euclídea, el termino “Euclídeo” es usado con motivo de distinguir estos espacios y los espacios curvos de geometría no Euclídea. Las operaciones mas típicas en geometría Euclídea puede ser representada con matrices de transformación si se usan sistemas de coordenadas homogéneos. Debido a esto una transformación T puede ser usada para diferentes motivos.

Usando conceptos básicos de álgebra lineal, una matriz 4×4 puede ser usada para expresar transformaciones lineales de un punto o un vector. Una transformación entonces será representada por los elementos de la matriz 4×4 . Ellas puede ser tambien usadas para realizar algunas transformaciones no lineales en espacios Euclídeos. Por esta razon, las matrices de transformación son altamente usadas en esta disciplina.

Generalmente una transformación es una función de puntos a puntos o vectores a vectores, por ejemplo:

$$p' = \mathbf{T}(p) \quad v' = \mathbf{T}(v) \quad (3.1)$$

Para transformar los puntos o vectores simplemente se tiene que realizar las apropiadas multiplicaciones de matrices. Esto también permite la composición de transformaciones, multiplicando las matrices de transformación.

Las transformaciones mas utilizadas comunmente son:

- **Rotación**, con el que se podrá rotar un punto o vector según un ángulo dado alrededor de un eje arbitrario o los ejes x , y o z .
- **Escalado**, para un punto o vector dado, esta transformación escalará los componentes x , y y z por un factor.
- **Translación**, únicamente afecta a puntos y podrá transladar las coordenadas x , y y z una cantidad.
- **Modelo**, esta matriz es útil para transformaciones del modelo localmente. This matrix is useful for transforming the model locally. De estar compuesto por un conjunto de matrices de rotación, escalado o translación. Cuando sea aplicada esta matriz a todos los puntos, los tendremos en coordenadas del mundo.
- **Vista**, inicialmente la cámara es localizada en el origen del espacio mundo. Para poder moverla alrededor del espacio mundo esta matriz es utilizada. Esta matriz tambien es denominada como “mirar a”. Luego de que esta transformation sea aplicada, tendremos los puntos en coordenadas de la cámara.
- **Proyección**, puesto que la escena tiene que ser proyectada tanto en una pantalla como en una imagen 2D, necesitamos otro proceso que convierta los puntos desde las coordenadas de la cámara a coordenadas homogéneas de modo que estas puedan ser proyectadas en pantalla.

Normalmente la composición de las matrices del modelo, vista y proyección es llamado *MVP* y es aplicada a cada punto que queremos dibujar.

Para ilustrar como las transformaciones son representadas por una matriz 4×4 , la ecuación 3.2 muestra la matriz genérica para la translación:

$$\mathbf{T}(\Delta x, \Delta y, \Delta z) = \begin{pmatrix} 1 & 0 & 0 & \Delta x \\ 0 & 1 & 0 & \Delta y \\ 0 & 0 & 1 & \Delta z \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (3.2)$$

Esta transformación es aplicada al punto $P = (x, y, z)$ de la siguiente forma:

$$\begin{pmatrix} 1 & 0 & 0 & \Delta x \\ 0 & 1 & 0 & \Delta y \\ 0 & 0 & 1 & \Delta z \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} x + \Delta x \\ y + \Delta y \\ z + \Delta z \\ 1 \end{pmatrix} \quad (3.3)$$

3.3. Camera model

Casi todo el mundo actualmente ha usado una cámara y conoce su propósito: se quiere obtener una imagen del mundo (normalmente presionando un botón) y la imagen es entonces almacenada en un *film*. Una de las cámaras mas simples es la cámara estenopeica. Este tipo de cámaras están compuestas por: una caja estanca, un pequeño orificio por donde entra la luz (estenopo) y un material fotosensible. De modo que cuando el pequeño orificio es descubierto la luz entra a través de el y alcanza la pieza de material fotosensible que está dispuesto al otro lado de la caja (ver Figura 3.2). Hoy en dia, las cámaras son mucho mas complejas de funcionamiento que esta simple cámara, pero es un buen comienzo para explicar como el funcionamiento de nuestra simulación.

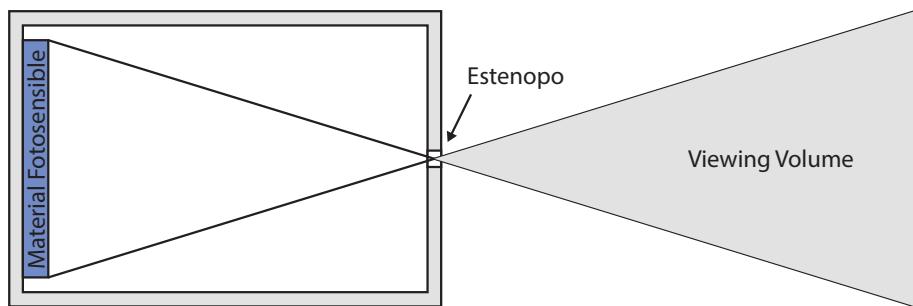


Figura 3.2: Funcionamiento de una cámara estenopeica.

Lo más importante es definir que parte de la escena será reflejada en el papel fotosensible. Conectando el estenopo con las aristas del papel se genera una doble pirámide que extiende dentro de la escena. Los objetos que no estean dentro de esta pirámide no quedarán

reflejados en la imagen del film. Como las cámaras de hoy en dia son mas complejas, se referirá a la región que puede ser reflejada en la imagen final como *viewing volume*.

Puesto que la imagen obtenida es la situada enfrente de el estenopo, este comunmente es también referido como *ojo*. Una cantidad de luz viajará del mundo al plano donde está situado el material fotosensible. Varios tests serán realizados por OpenGL dependiendo del tipo de cámara para establecer que puntos pueden ser vistos y tendrán que ser representados en el *frame-buffer*.

El modelo de la cámara usado en este proyecto soporta diferentes configuraciones. Primero se explicará como estos parámetros afectan al modelo de la cámara, luego se explicará como el modelo interacciona con el proceso de *ray tracing*.

3.3.1. Parámetros de la Cámara

El primer parámetro que el modelo de la cámara soporta es la **posición**. Este parámetro establece donde la cámara tendrá que estar situada en el espacio de coordenadas del mundo (x,y,z) .

El siguiente parámetro soportado es el de la rotación. Este parámetro es un punto en el mundo hacia donde la cámara estará mirando. También necesitamos saber cómo orientar la cámara a lo largo de la dirección de visión implícita por los dos primeros parámetros .

También es necesario conocer como está orientada la cámara con respecto a la dirección en la que se está mirando implicado por los 2 primeros parámetros. El parámetro *up vector* nos da esa orientación.

Podremos ver como estos parámetros están organizados en el espacio de la cámara en .

Otros parámetros importantes en el modelo de la cámara son los planos de *clipping*, que nos darán el rango a lo largo del eje z que dejará que los objetos dentro de el sean renderizados. **Near** establece la posición del plano cercano y **far** indica la posición del plano de *clipping* lejano.

El último parámetro que la cámara soporta es el **campo de visión**. El ángulo de visión describe la extensión angular de la escena capturada por la cámara horizontal y verticalmente .

Los planos de *clipping* y los ángulos de visión define el *viewing volume* en el modelo, también conocido como *viewing frustum*.

El usuario también tiene la opción de usar entre una cámara con proyección **perspectiva** o **ortográfica** dependiendo de las necesidades. En arquitectura por ejemplo se suele

preferir una cámara ortográfica, mientras que un usuario normal suele preferir una en perspectiva puesto que el campo de visión da un resultado más natural. (ver Figura 3.3)

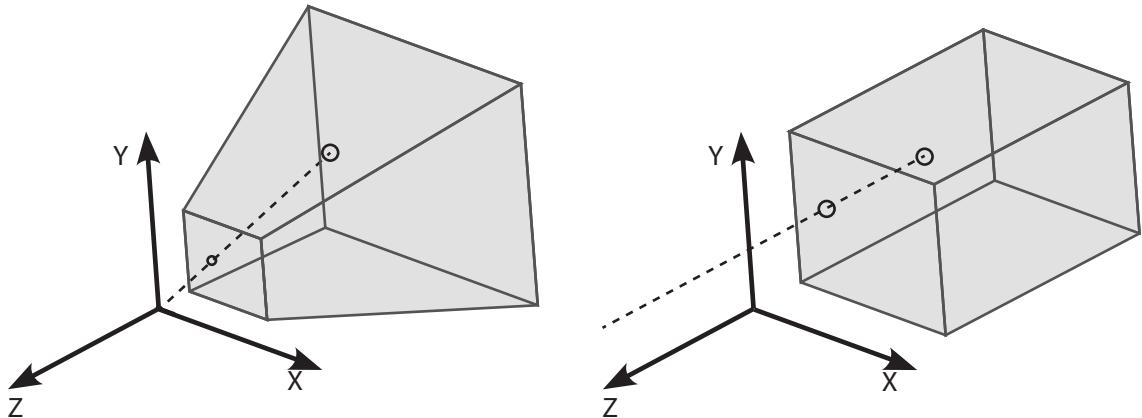


Figura 3.3: Representación de las diferencias entre cámara en perspectiva (izquierda) y ortográfica (derecha).

La proyección ortográfica es mas comunmente utilizada en ingeniería puesto que las dimensiones son mostradas de forma inequívoca, si hay alguna linea de unidad en cualquier lado de la escena, esta se mostrara con la misma longitud en cualquier lado. También cualquier par de lineas que sean paralelas lo serán también en la realidad. Con la proyección en perspectiva, lineas de identico tamaño en la vida real podrán aparecer de longitud diferente debido al *escorzo*.¹. Esto hace que sea difícil juzgar las dimensiones relativas de un objeto cuando está lejano.

3.3.2. Funcionamiento interno de la cámara

Una vez que se tienen los parámetros que se necesitan para localizar la cámara en escena, usaremos una transformación *mirar a* para este propósito. Usando esta transformación como matriz de vista nos dará una transformación entre coordenadas del mundo y la cámara como se mencionó anteriormente. La matriz es construida usando la posición (p), orientación (o) y up vector (u).

Primero, un vector dirección f es construido:

$$f = \frac{p - o}{\|p - o\|} \quad (3.4)$$

Después, el vector *right r*:

¹Los objetos se vuelven más pequeños en la imágen en cuanto estén situados en un plano más lejano.

$$r = f \times u \quad (3.5)$$

Luego, un nuevo vector $up\ u'$ es calculado en función de la referencia de la cámara:

$$u' = r \times f \quad (3.6)$$

Con éstos, una matriz de rotación se puede construir lo que representa una reorientación en la base ortonormal de nueva creación :

$$R = \begin{bmatrix} r_x & u'_x & f_x & 0 \\ r_y & u'_y & f_y & 0 \\ r_z & u'_z & f_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.7)$$

Por último, para transformar objetos en el marco de la cámara; no sólo todo tiene que ser orientada correctamente, sino que también tiene que ser traducido desde el origen hasta la posición de la cámara. Esto se logra mediante la concatenación de la matriz de rotación R con la matriz correspondiente:

$$V = \begin{bmatrix} r_x & u'_x & f_x & -p_x \\ r_y & u'_y & f_y & -p_y \\ r_z & u'_z & f_z & -p_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.8)$$

La ecuación 3.8 es la resultante matriz *mirar a* V que será usada para transformar los puntos del espacio mundo a coordenadas de la cámara.

La siguiente fase dependerá del tipo de cámara que el usuario elija, ortogonal o perspectiva. En función de esta elección una matriz de proyección será escogida.

Si es escogida una cámara perspectiva, tendremos que tener en cuenta el efecto del escorzo mencionado anteriormente. Esta proyección no conserva distancias, ángulos, y paralelismo entre líneas. x e y serán modificadas de acuerdo al correspondiente valor de z , que establece como de cerca a la cámara está un punto. Para obtener este efecto, una matriz perspectiva será utilizada. Una matriz comúnmente utilizada es la matriz *frustum*. En ella el espacio homogéneo obtiene la forma de una pirámide truncada. Los parámetros usados para la construcción de esta matriz serían el **izquidara** (l), **derecha** (r), **arriba** (t), **abajo** (b), plano **near** (n) y **far** (f) del *frustum*.

t es obtenida de la siguiente forma usando los parámetros de la cámara del **campo de visión** (*fov*) y el plano *near* (*n*):

$$t = n \cdot \tan(fov/2) \quad (3.9)$$

Luego, r es obtenida usando el ancho (*w*) y alto (*h*) de la ventana:

$$r = t \cdot w/h \quad (3.10)$$

Puesto que el *frustum* es simétrico, l y b serán lo mismo que r y t pero de signo opuesto. Una vez se tienen todos los parámetros se podrá construir las matrices de perspectiva u otrográfica:

$$P = \begin{bmatrix} \frac{2 \cdot n}{r - l} & 0 & \frac{r + l}{r - l} & 0 \\ 0 & \frac{2 \cdot n}{t - b} & \frac{t + b}{t - b} & 0 \\ 0 & 0 & \frac{n + f}{n - f} & \frac{2 \cdot n \cdot f}{n - f} \\ 0 & 0 & -1 & 0 \end{bmatrix} \quad (3.11)$$

En contraste, si la cámara requerida es la ortográfica, In contrast, if a orthographic camera is required, una matriz algo más simple se puede usar. Los puntos serán proyectados al plano *near*. Este tipo de transformaciones no da el efecto de escorzo, manteniendo paralelas y manteniendo la distancia relativa entre objetos. Esta transformación mantiene *x* e *y* invariables, pero cambia el valor de *z* al plano *near*. Los parámetros para la construcción de esta matriz son también **izquierda** (*l*), **derecha** (*r*), **arriba** (*t*), **abajo** (*b*) y plano **near** (*n*) del *frustum*:

$$P = \begin{bmatrix} \frac{2}{r - l} & 0 & 0 & \frac{l + r}{l - r} \\ 0 & \frac{2}{t - b} & 0 & \frac{b + t}{b - t} \\ 0 & 0 & \frac{2}{n - f} & \frac{n + f}{f - n} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.12)$$

Una de estas dos matrices (3.11 or 3.12) representarán la matriz de proyección \mathbf{P} .

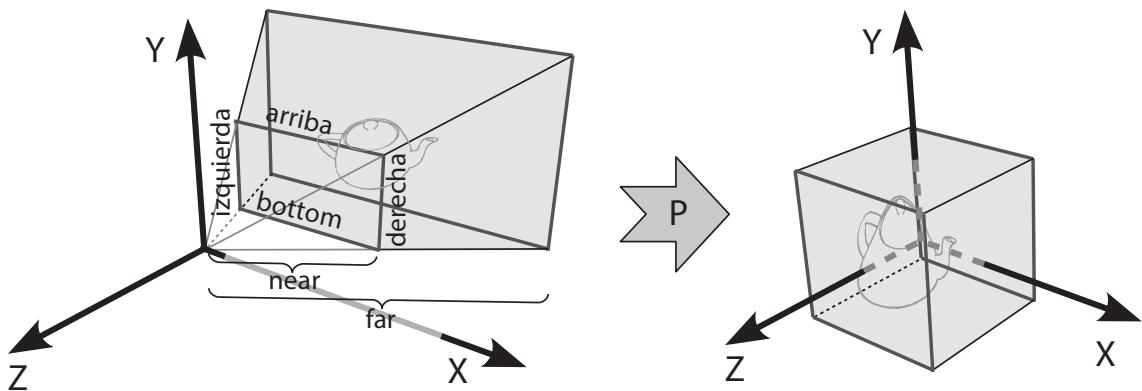


Figura 3.4: La matriz P se encarga de transformar la pirámide de visualización en un cubo unitario.

La composición de estas matrices producirá la matriz MVP. Esta es la expresión matemática usada para transformar cada punto:

$$\begin{aligned} v' &= \mathbf{P} \cdot \mathbf{V} \cdot \mathbf{M} \cdot v \\ &= \mathbf{MVP} \cdot v \end{aligned} \tag{3.13}$$

Estos son los paso necesarios para la creación de un frame, pero puesto que le visualizador es interactivo, tendremos que repetir este proceso varias veces por segundo.

Capítulo 4

Visualización de nubes de puntos con GPU

Este capítulo se presenta como un resumen a los diferentes métodos de render basados en GPU. La sección 4.1 empieza mostrando los diferentes algoritmos que existen para la rasterización de splats. La sección 4.2 se centra en el blending de los diferentes splats, rasterizados en el paso previo, teniendo en cuenta la implicación que toman diferentes aspectos como la iluminación.

4.1. Técnicas de rasterización

El primer paso en el render de nubes de puntos, es determinar según la proyección de los splats que píxeles en el framebuffer serán cubiertos por ellos. Puesto que no existe soporte en las librerías actuales para representar este tipo de primitivas, los splats tendrán que ser representados por otras como puntos o triángulos. De entre esas dos opciones, se considera como más eficiente los puntos por delante de los triángulos, ya que por cada splat únicamente hay que almacenar la posición de un vértice en lugar de tres. Finalmente para renderizar la nube, habrá que dibujar el *VBO*, pasándole el tipo *GL_POINTS*

```
glDrawArrays(GL_POINTS, &splats[0], splats.size());
```

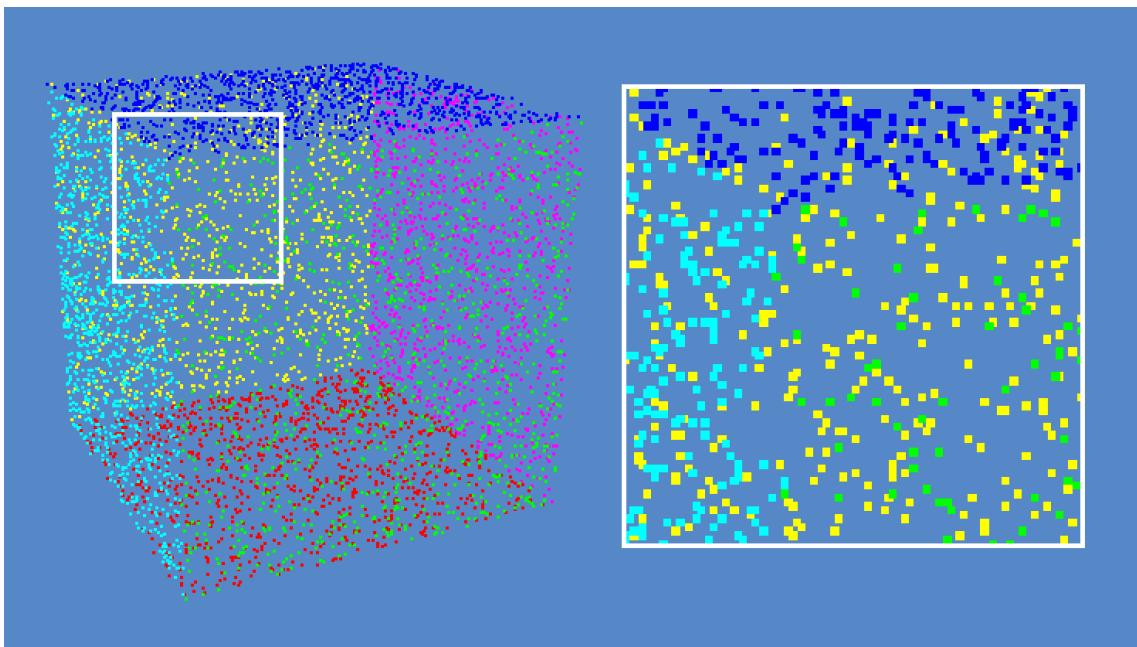


Figura 4.1: Todos los puntos de la nube son renderizados con un cuadrado del mismo tamaño para toda la nube.

4.1.1. Sized-Fixed Points

De una forma un poco primitiva y puesto que un punto a priori puede no tener un tamaño definido, a la hora de mostrar una representación de la nube, se podría dibujar cada coordenada del conjunto de puntos, con un cuadrado de un tamaño definido. Para ello basta con activar en el programa.

```
glEnable(GL_PROGRAM_POINT_SIZE);
```

Mientras que en el *vertex shader*, se define con *gl_PointSize* el tamaño en pixels que se pretende que ocupe el punto al ser rasterizado. Este cuadrado tendrá como centro las coordenadas del punto **p**.

Como resultado, se obtiene una representación en la que cada punto de la nube se renderizará con el tamaño descrito.

4.1.2. Image-aligned Squares

En el anterior paso, el tamaño de la representación que cada punto tomaba en el *frame-buffer* era de alguna forma independiente a la cámara y al modelo.

De esta manera, si en lugar de entender el tamaño del punto como los pixeles que va a ocupar en la imagen final, se interpreta como una medida dentro del modelo, que podría ser el radio \mathbf{r} del splat, al representarse, estos puntos tendrían que ser coherentes con su proyección.

Así que el tamaño en el *screen-space* del punto OpenGL proyectado, tiene que ser ajustado en el *vertex shader*. El tamaño [1] se aproxima haciendo el escorzo (*perspective foreshort*) del radio r del splat usando el *depth value* del centro \mathbf{p} que ha de estar en coordenadas de la cámara.

$$gl_PointSize = 2r \cdot \frac{n}{p_z} \cdot \frac{g}{t - b'} \quad (4.1)$$

Donde n , t y b son los parametros *near/top/bottom* del *view frustum* y h denota la altura (en pixels) del *viewport*. En esta fórmula el término n/z corresponde a la proyección en el plano *near*, mientras que $f/(t - b)$ escala el resultado desde el plano *near* a coordenadas de la imagen.

Todos los pixels, generados de un punto, tienen el mismo *depth value* lo que imposibilita hacer uso de las técnicas blending en zonas que visualmente se estén superponiendo. (ver sección 4.2)

4.1.3. Affinely Projected Point Sprites

Una mejor aproximación para representar superficies afines a nubes de puntos, es con discos orientados según un vector normal \mathbf{n} , este método fue presentado en *Bosch and Kobbelt*.

Para ello se tendría que ajustar el tamaño del punto en el *vertex shader* de la misma forma que en el paso anterior 4.1.2. Pero añadiendo a mayores en el *fragment shader* alguna manera de determinar si este se corresponde con un fragmento de la proyección de un disco orientado con centro en la posición \mathbf{p} , de radio \mathbf{r} y con la normal \mathbf{n} .

Para cada uno de los pixeles $(x, y) \in [-r, r]$, se puede calcular un valor de desplazamiento en profundidad δz como una función lineal dependiente del vector en coordenadas de la cámara $\mathbf{n} = (n_x, n_y, n_z)^T$:

$$\delta z = -\frac{n_x}{n_z} \cdot x - \frac{n_y}{n_z} \cdot y \quad (4.2)$$

Este desplazamiento es usado para calcular la distancia 3D desde el centro del punto \mathbf{p} , de manera que el fragmento (x, y) corresponde al punto si $\|(x, y, \delta z)\| \leq r$. Los que no cumplan esta condición pueden ser fácilmente descartados usando el comando *discard*.

La variable vector *gl_PointCoord* del *fragment shader* contiene las coordenadas bidimensionales que indican donde dentro de un punto OpenGL el fragmento está situado, con valores desde el 0,0 al 1,0 estos han de desplazarse a un rango $[-0,5, 0,5]$ para usarse en el cálculo de δz .

Una posible implementación de este algoritmo en *GLSL* sería:

```
vec3 test;

test.x = gl_PointCoord.x - 0.5;
test.y = gl_PointCoord.y - 0.5;
test.z = -(normal.x/normal.z) * test.x - (normal.y/normal.z) * test.y;

if (length(test) > 0.5)
    discard;
```

Uno de los inconvenientes usando Image-aligned Squares 4.1.2 era que el valor de profundidad era constante por splat. Pero δz puede ser usado también para corregir esto. Partiendo de un valor de profundidad en espacio de la cámara $z' = p_z + \delta z$ y el frustum el valor de profundidad del *zbuffer* (x, y) puede ser modificado

$$zbuffer(x, y) = \frac{1}{z'} \cdot \frac{fn}{f - n} + \frac{f}{f - n} \quad (4.3)$$

En comparación con la propuesta de *Image-aligned Squares*, este método ofrece una mejor aproximación, especialmente en los contornos del objeto. Sin embargo el valor de profundidad δz es solo una aproximación, ya que se asume una proyección paralela (4.2), causando errores que probocan que los discos se vuelvan demasiado finos cuando estos son mirados bajo ángulos planos, lo que puede resultar en huecos en la imagen renderizada (ver Figura 4.2).

4.1.4. Perspective Correct Rasterization

El método de *Affinely Projected Point Sprites* (4.1.3) causaba errores debido a la proyección paralela que en esta se asumía. Puesto que la proyección puede o no ser paralela, se tendrá que usar una aproximación mas precisa. Esta técnica fué la propuesta por *Bosch et al.*

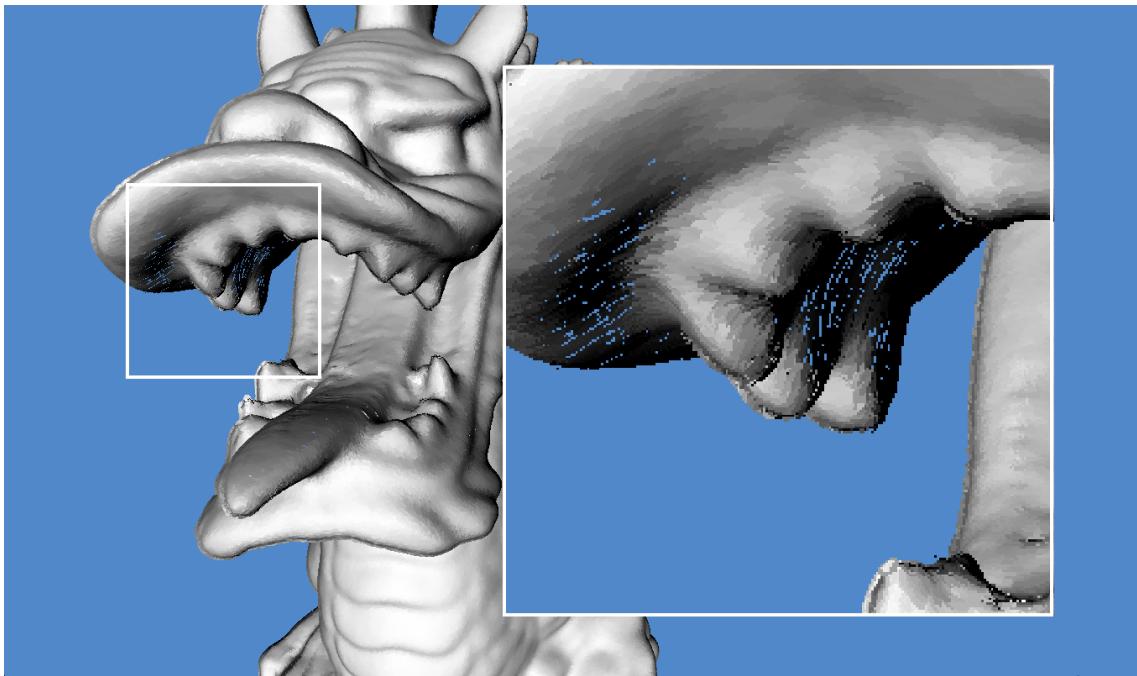


Figura 4.2: Dragon, modelo de 845,281 puntos renderizado mediante Affinely Projected Point Sprites. Este método causa huecos en ángulos extremos.

La idea principal de este procedimiento parte en determinar el punto 3D correspondiente al pixel 2D mediante un *ray casting* local.

Fijándose en el *pipeline* de transformaciones de OpenGL el primer paso para calcular \mathbf{q} es invertir la transformación *window-to-viewport*, de esta manera se mapeará el pixel (x,y) a un punto 3D \mathbf{q}_n en el plano *near*.

$$\mathbf{q}_n = \begin{pmatrix} x \cdot \frac{r-l}{w} - \frac{r-l}{2} \\ y \cdot \frac{t-b}{h} - \frac{t-b}{2} \\ -n \end{pmatrix} \quad (4.4)$$

Donde x/y son las coordenadas relativas a la ventana *gl_FragCoord.xy*, $b/t/l/r$ son los parámetros *bottom/top/left/right* del *viewing frustum*, y w/h denotan la *altura/anchura* del *viewport* (ver Figura 4.3).

Se proyectará un rayo desde el origen (el ojo de la cámara) a través de \mathbf{q}_n y la intersección con el plano del *splat* \mathbf{p} dará como resultado el punto \mathbf{q} . En caso de que $\mathbf{q}_n \cdot n = 0$ se determinará que el el ángulo el por lo que se podrá descartar el *splat*.

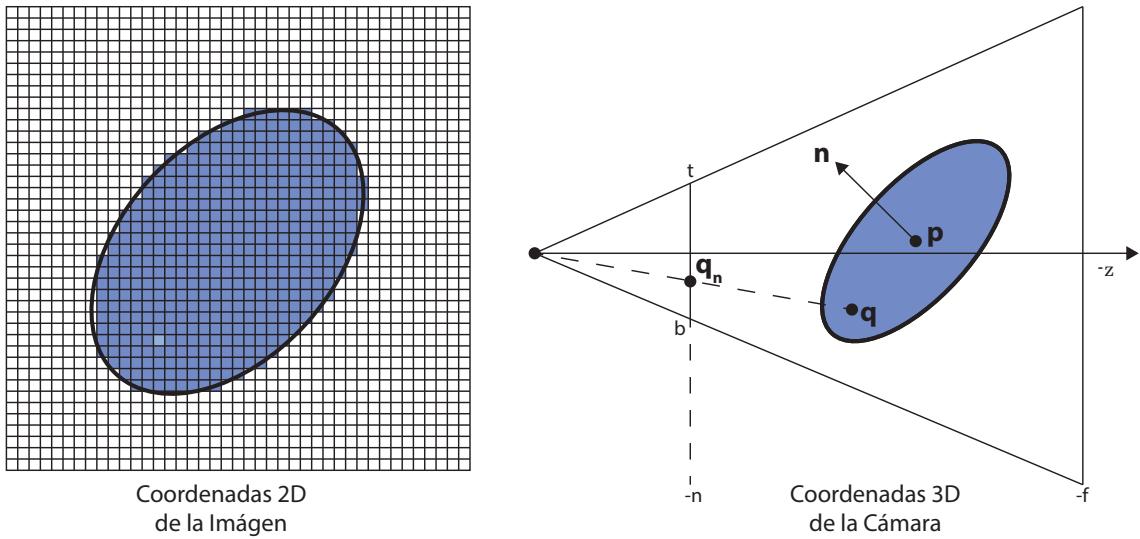


Figura 4.3: Proyección y Ray Casting.

$$q = q_n * \frac{p \cdot n}{q_n \cdot n} \quad (4.5)$$

Finalmente el pixel será descartado si $\|q - p\| > r$. Por otro lado el valor de profundidad del *z-buffer* puede ser ajustado insertando q_z como valor de z' en la ecuación (4.3).

4.2. Técnicas de Blending

Luego de determinar los pixels que son cubiertos por la proyección de los *splats* en el anterior paso de rasterización, el siguiente implica iluminación y en general el acabado estético de la superficie *shading*.

4.2.1. Flat Shading

Puesto que cada punto tiene asociado un vector normal **n**, para unas ciertas propiedades de material y reflectancia una iluminación local puede ser evaluada por *splat*. Como resultado se obtendrá una constante de color c_i para cada *splat*, similar al *flat shading* en modelos poligonales. Ya que los puntos tienden a intersecar unos con otros, las representaciones mediante *flat shading* conllevan a discontinuidades en el color.

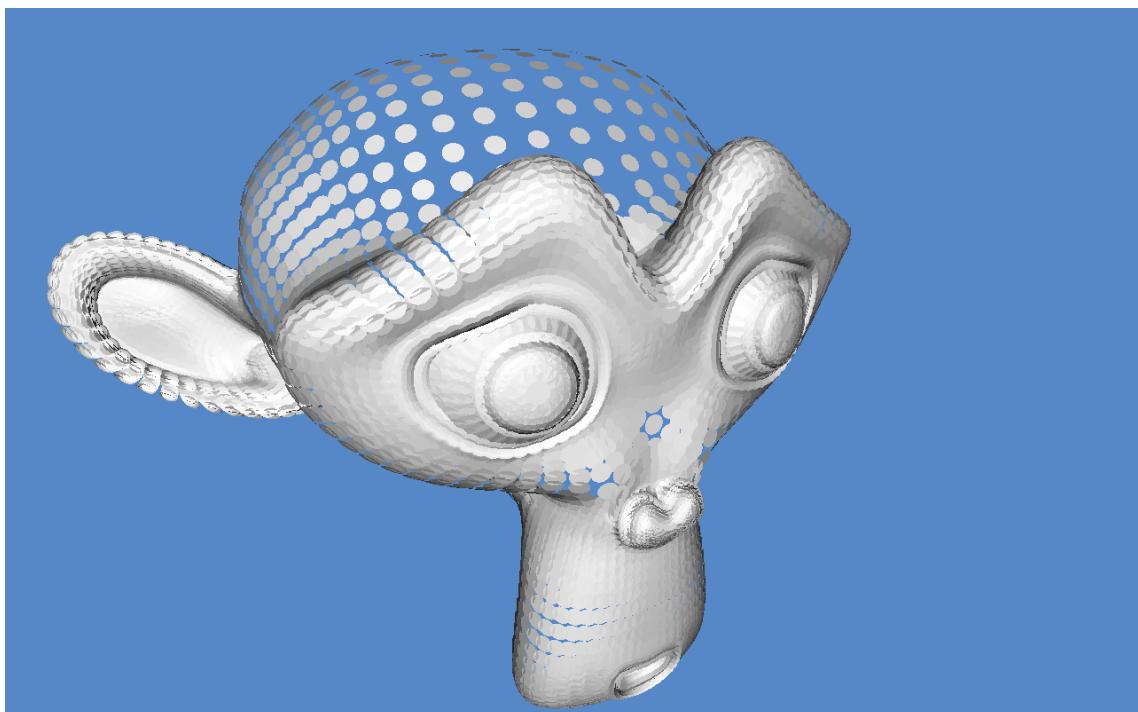


Figura 4.4: Suzanne, modelo de 7,958 puntos renderizado mediante Perspective Correct Rasterization.



Figura 4.5: Lucy, modelo de 397,664 puntos. Renderizado mediante *Flat Shading* (izquierda), *Gouraud Shading* (centro), y *Phong Shading* (derecha).

4.2.2. Gouraud Shading

Para conseguir un render mas fino, las discontinuidades en el color probocadas por el *shading* pueden ser emborronadas mediante *blending* de los valores de color \mathbf{c}_i de los puntos que se están superponiendo. Puesto que los cálculos de iluminación todavía son computados por *splat*, este tipo de *blending* del color, conceptualmente corresponde al *Gouraud shading* de modelos poligonales.

Para implementar este tipo de *blending* cada punto de la nube irá asociado con una función, de modo que para cada *splat* sus fragmentos irán teniendo menor peso \mathbf{r} cuanto mas alejados de su centro estén, teniendo así menor impacto en la mezcla final de color $\mathbf{c}(\mathbf{x}, \mathbf{y})$. De manera que en la reconstrucción final del *framebuffer*, el color de la posición (x, y) vendrá dado por el peso promedio de todos los fragmentos i que intenten cubrirlo.

$$\mathbf{c}(x, y) = \frac{\sum_i r_i(x, y) \mathbf{c}_i}{\sum_i r_i(x, y)} \quad (4.6)$$

Esta media puede ser implementada en OpenGL usando 2 pasos de render. Primero, el *alpha-blending* es configurado con funciones separadas de *blend* para las componentes RGB y alpha *EXT_blend_func_separate* con

```
glEnable(GL_BLEND);
glBlendFuncSeparateEXT(GL_SRC_ALPHA, GL_ONE, GL_ONE, GL_ONE);
```

Así los *splats* irán acumulando los valores de color y peso en las componentes RGB y alpha del *framebuffer* como $(\sum_i r_i c_i, \sum_i r_i)$. Finalmente tendrá que normalizarse la componente *RGB* de cada pixel, dividiéndose por su componente *alpha*. Esto puede ser conseguido enviando el resultado del anterior pase a un siguiente como una textura, renderizando un rectángulo de tamaño de la ventana y mapeando esta textura en el. Este truco conocido como *render-to-texture*, envía de nuevo cada pixel a través del *pipeline* de OpenGL, de modo que un simple *fragment shader* puede realizar esta normalización. Esta acumulación debe de ser realizada usando *buffers* con precisión de punto flotante de 16 bits, con motivo de evitar saturaciones o artefactos.

Para restringir el *blend* únicamente a la superposición de los *splats* vecinos de la misma superficie. Es necesario un tercer pase, denominado de visibilidad. Para cada *frame*, el pase de visibilidad primero renderiza la nube únicamente en el *depth buffer*. Luego en el pase de *blending* se renderiza la nube de nuevo, pero esta vez calculando iluminación y acumulando el valor resultante de color usando *additive alpha-blending* como lo descrito anteriormente. En este pase no se debe de actualizar el *z-buffer* calculado en el pase de

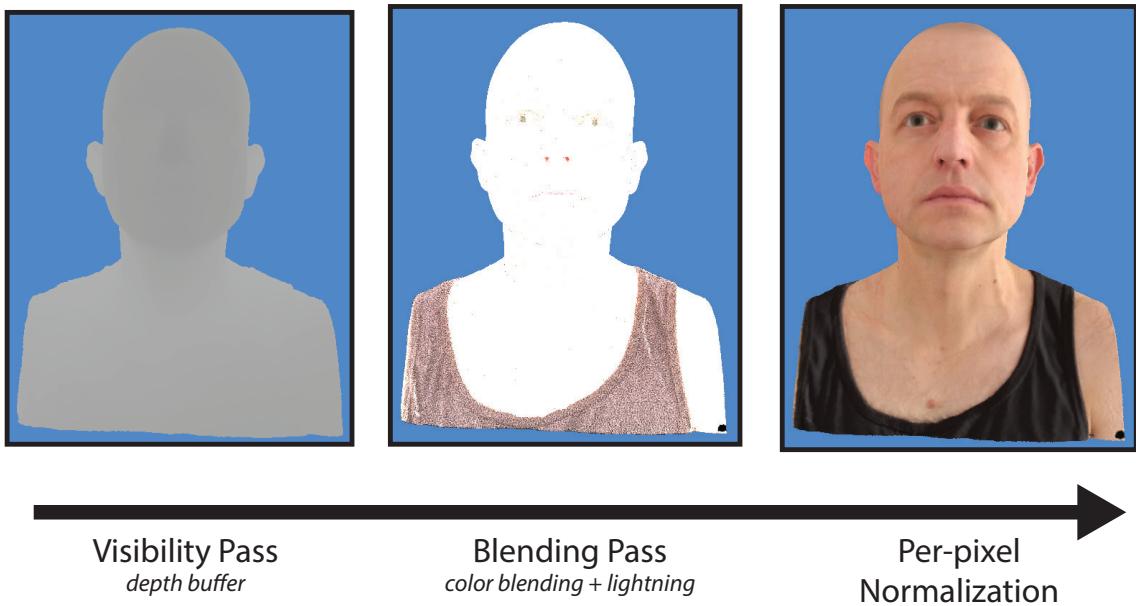


Figura 4.6: Los tres pasos, *visibility*, *blending*, y *normalization*.

visibilidad, pero si se añade un desplazamiento ϵ a todos los valores de profundidad, lo que probocará que todos los fragmentos dentro de una distancia ϵ sean mezclados. El pase final de normalización realiza la necesaria division por el componente *alpha* como se describió anteriormente. El resultado de los tres pasos de render es representado en la Figura 4.6.

El *pseudocode* sería el siguiente

```

// visibility pass
glDepthMask(GL_TRUE);
glDisable(GL_BLEND);
bind_visibility_shaders();
glDrawArrays(GL_POINTS, &cloud[0], cloud.size());

// blending pass
glDepthMask(GL_FALSE);
 glEnable(GL_BLEND);
bind_blending_shaders();
glDrawArrays(GL_POINTS, &cloud[0], cloud.size());

// normalization pass
glCopyTexSubImage2D(GL_TEXTURE_RECTANGLE_ARB, 0, 0, 0, 0, w, h);
bind_normalization_shaders();
draw_rectangle();

```

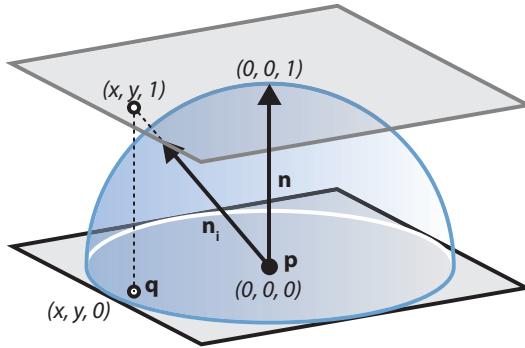


Figura 4.7: Los vectores n_i son representados como puntos homogéneos $(x,y,1)$ en un plano desplazado tangente.

4.2.3. Phong Shading

Cuando se comparan los resultados de las diferentes técnicas en la figura 4.5, *Gouraud* realmente elimina las indeseadas discontinuidades de color provocadas en el *Flat Shading*, pero también añade cierto emborroneo a la imagen notablemente. Para renderizado de polígonos es bien conocido que *Phong* es superior que *Gouraud Shading*. En lugar de calcular la iluminación por cada vértice y linealmente interpolar el color resultante dentro de los triángulos, *Phong* interpola las normales en los vértices, seguido de una iluminación por píxel basada en la resultante función definida a trozos.

Pero ya que en el caso de las nubes de puntos falta la relación de conectividad, no es posible la interpolación de las normales de los puntos vecinales, de modo que el campo de normales deberá ser construido de otra forma.

El método de *Phong* de Botsch asigna explícitamente un campo de normales lineal $\mathbf{n}_i(x, y)$ para cada *splat* en lugar de mantener su normal asociada constante. Durante la rasterización, la normal es evaluada para cada pixel basándose en sus parámetros locales (x, y) , y el vector resultante $\mathbf{n}_i(x, y)$ es usado para computar la iluminación. Ya que los valores de color resultantes siguen teniendo problemas de discontinuidad, estos son acumulados y mezclados usando el mismo sistema de 3 pasos de *Gouraud*. La única que queda por conocer entonces es saber cómo computar el campo de normales lineal $\mathbf{n}_i(x, y)$. Cada normal será representada por un punto (x, y) en el plano tangente con distancia 1, similar a las coordenadas homogéneas (Figura 4.7).

$$\mathbf{n}_i = \frac{(q + n) - p}{\|(q + n) - p\|} \quad (4.7)$$

Donde \mathbf{q} es el punto 3D del *splat* calculado al igual que en el método de *Perspective*

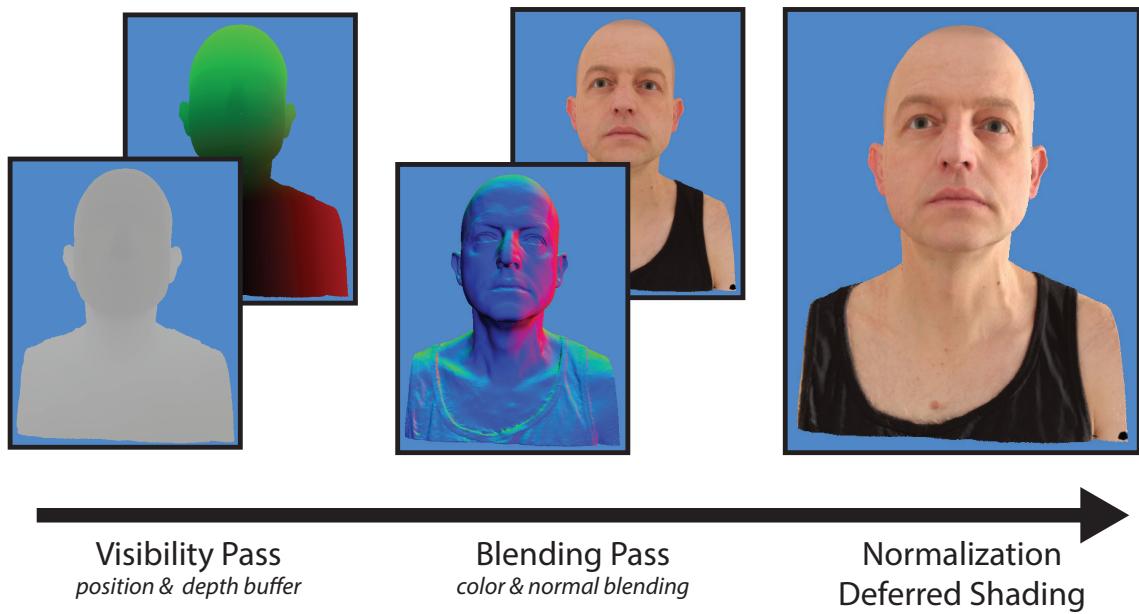


Figura 4.8: El renderizado mediante *Deferred Shading* acumula atributos como color y normales, seguido por una normalización y un shading pass.

Correct Rasterization (Ecuación 4.5), \mathbf{n} es la normal del *splat* y \mathbf{p} es el centro de este en coordenadas de la cámara.

4.2.4. Deferred Shading

Básicamente existen dos formas para generar interpolaciones suavizadas de vectores de normales. La primera es *Phong* que como se comentó en el anterior apartado, utiliza funciones para calcular el campo de normales de un *splat*. Una segunda opción pasaría por la denominada como *deferred* en la que en lugar de mezclar en un misma fase color y normales, estas se calculan de forma separada.

Recientes generaciones de *GPU* contienen todo lo necesario para implementar esta aproximación, que gracias al denominado *multiple render targets* (*ARB_DRAW_BUFFERS*) permite que en un simple pase de *render* se puedan sacar diferentes resultados a diferentes *buffers*.

Primeramente en el pase de visibilidad (ver Figura 4.8, izquierda), se hará una modificación para poder sacar en un target a parte la posición 3D del fragmento, \mathbf{q} 4.5 ya que será necesario para el pase final. Luego, en el pase de acumulación y blend (ver Figura 4.8, centro) se usarán 2 targets adicionales para sacar la salida referente al color y de las normales de manera separada, usando el mismo proceso de acumulación que en *Gouraud*. Finalmente estos 3 *buffers* serán enviados de nuevo al *openline* de OpenGL, usando el

método de *render to texture* comentado previamente, con motivo de su normalización.

Este último pase (ver Figura 4.8, derecha), corresponde a la fase de normalización de *Gouraud* pero en el que a mayores se van a realizar los cálculos pertinentes para la iluminación. Por cada pixel, una media de la normal y el color pueden ser obtenidos, que junto a la posición obtenida en la primera fase de visualización puede ser fácilmente calcular el color final iluminado.

Es importante darse de cuenta de que en este método los cálculos derivados de la iluminación son computados una única vez por pixel, en contraste con las aproximaciones que no son *deferred* que incorporan este proceso en el momento de la rasterización, consiguiendo el color final del pixel a partir de la mezcla de muchos pixels de diferentes *splats* iluminados previamente. Dependiendo del número de *splats* superpuestos, esto puede llegar a multiplicar el numero de cálculos derivados de la iluminación.

Además a esto, *deferred shading* añade también una clara separación entre la fase de rasterización y de iluminación.

Capítulo 5

Resultados y rendimiento

Proin sit amet est. Nunc vel sem sed ligula accumsan cursus. Nunc tincidunt purus non elit accumsan vestibulum. Proin ultrices metus non nisl. Duis feugiat enim et sapien. Suspendisse lectus. Aliquam erat volutpat. Vestibulum ullamcorper. Nullam augue urna, pharetra nec, feugiat at, pellentesque a, nibh. Curabitur nulla.

Capítulo 6

Conclusiones

Proin sit amet est. Nunc vel sem sed ligula accumsan cursus. Nunc tincidunt purus non elit accumsan vestibulum. Proin ultrices metus non nisl. Duis feugiat enim et sapien. Suspendisse lectus. Aliquam erat volutpat. Vestibulum ullamcorper. Nullam augue urna, pharetra nec, feugiat at, pellentesque a, nibh. Curabitur nulla.

Bibliografía

- [1] Matthias Zwicker, Jussi Räsänen, Mario Botsch, Carsten Dachsbacher, and Mark Pauly. Perspective accurate splatting. *Proceedings of Graphics Interface*, pages 247–254, 2004.