

Título

subtitulo

JAVIER PINA DE PAZ

5 de julio de 2018

Dedico este TFG a mi estupidez por
matricularme antes de tiempo :D

Índice general

1. Capítulo introductorio	1
1.1. La visión como herramienta	1
1.2. Concepto de nube de puntos	4
1.3. Adquisición de información: Sensores y evolución	5
1.4. Procesamiento software de información	6
1.5. Descripción de herramientas	6
1.5.1. Herramientas hardware	6
1.5.2. Herramientas software	6
1.6. objetivos	6

1

Capítulo introductorio

“ Me cago en Dios”

– Javier, *Todos los días*

1.1. La visión como herramienta

El ser humano es capaz de interactuar con el mundo que le rodea de diversas maneras ya sean creativas o incluso productivas. Esto se hace con el uso de los sentidos; Tacto, gusto, oído, olfato y por último pero no menos importante la visión. Este sentido nos permite recoger gran cantidad de información del mundo que nos rodea, procesarla y actuar en consecuencia. Es entonces el ojo humano una herramienta de gran valor y elevada precisión que poco tiene que envidiar al órgano semejante en el resto de especies en todo el reino animal. Lejos de realizar un estudio anatómico del sentido de la vista en el ser humano, se disponen a continuación ciertas características de interés que permiten apreciar toda su potencia.

Una de las primeras cuestiones que se vienen a la mente cuando se habla de la "simple vista"^{es} su resolución, es decir, cuál es la mínima distancia que es capaz de distinguir. se puede estudiar mediante la resolución angular que para el ser humano se encuentra entorno a 1' o 2' o lo que es lo mismo un intervalo de 0,02° a 0,03°. Dicho de otra forma, un ojo humano sano y correctamente desarrollado puede distinguir objetos de entre 30cm y 60cm a 1km de distancia. Por ejemplo, El ojo humano podría servir para detectar dos balones de playa de un diámetro aproximado de 45cm hasta 1km de distancia como máximo. Otra manera de comprender la potencia de esta resolución sería posible equiparándola con la potencia de una cámara digital, para el caso, varios cientos de megapíxeles. Véase como ejemplo una cámara Canon de 50,6 Megapíxeles disponible a más de 3500 euros *comparar vision humana y MPíxeles*

El campo de visión es también un aspecto de gran importancia ya que determina la cantidad de información que los ojos pueden recibir en un

instante dado. El cono de visión "queda entorno a unos 130° en vertical y 160° en horizontal

Por último, se considera la velocidad de enfoque o acomodación del ojo humano

Se ha planteado entonces una situación en la que hay un pequeño problema, un instrumento de gran precisión y prolongada durabilidad embarcado en un ser vivo falible y en ocasiones errático. No sólo esto sino que nunca se llega a dominar por completo el manejo de éste órgano así como su uso en conjunto a las demás partes del cuerpo y siempre todo ello limitado por la velocidad de respuesta que puede proporcionar el sistema nervioso. Por ejemplo, ...

Teniendo en cuenta una posible definición de tecnología

"La aplicación de conocimientos científicos para la resolución de problemas mediante el diseño y creación de bienes y servicios"

junto a su etimología, pues se trata de una palabra de origen griego formada por (arte, técnica u oficio) y (el estudio de algo) es fácil apreciar que la evolución y desarrollo del ser humano durante toda su historia han venido ligadas a un progreso tecnológico incesante, vivo y capaz de abrirse paso incluso en las épocas más oscuras de la historia de la humanidad. *línea temporal tecnología*

Una vez descritas brevemente la potencia, capacidad y posibilidades que aporta la visión en el ser humano, se puede ir más allá, subir un nivel y dotar a entes ajenos al hombre de la capacidad de obtener y procesar información visual. Este concepto existe desde finales de la década de los sesenta y puede introducirse como visión artificial o visión por computador:

"Disciplina científica que incluye métodos para adquirir, procesar, analizar y comprender las imágenes del mundo real con el fin de producir información numérica o simbólica para que puedan ser tratados por un computador"

Esta forma de trabajar con información visual es posible debido a la puesta en conjunto de diferentes campos como la geometría, física o estadística y demás herramientas que se explicarán más adelante.

Bien es cierto que se plantea un problema ya que la forma en que el ojo humano percibe el mundo no es la misma en la que lo hace una máquina pues se pueden establecer sus diferencias como las mismas que hay entre una señal analógica y otra digital, respectivamente.

Véase en primer lugar la mencionada diferencia respecto a las señales:

- Señal analógica

Se trata de la representación de una magnitud física tal y como se percibe del entorno o se genera con algún instrumento y puede verse como una función matemática continua. La mayoría de las señales que se perciben son analógicas como ejemplo la intensidad de corriente

eléctrica, la temperatura, el sonido, presión y energía. De este modo existen señales analógicas periódicas (caracterizadas por amplitud y frecuencia) no periódicas (toman cualquier valor independientemente del tiempo) *ejemplo de señales matlab automatica*

Señal digital

Existe una extendida confusión en lo referente a la diferencia entre señal digital y señal discretizada. Se parte de la siguiente señal analógica y sus características en amplitud y frecuencia:

A partir de la señal anterior se llega a una señal discretizada tomando valores equidistantes en el tiempo:

Se llega entonces a una señal digital codificando cada uno de los valores discretos de la señal representada anteriormente. De este modo se tiene un número determinado de valores fijos pertenecientes a un conjunto y no intervalos.

Se puede mostrar el funcionamiento básico de un convertidor Analógico/digital para comprender la diferencia entre estos dos conceptos: *apuntes micros convertidor y diagrama*

Llevando este concepto al campo de los sistemas digitales, tal y como puede ser un ordenador, se llega al uso de la lógica de dos estados o binaria en la cual existen los estados alto H o 1 y bajo L o 0 en el caso de lógica positiva y H o 0 junto a L o 1 para la lógica negativa. Un ejemplo de señal digital para la lógica de dos estados:

señal digital

A pesar de parecer señales discontinuas, en realidad existen transiciones continuas de un estado a otro llamadas flancos de subida o bajada. Se puede apreciar a continuación la misma señal que se ha mostrado en el ejemplo anterior pero en este caso a una escala diferente para poder apreciar la mencionada transición.

señal digital a escala mayor

Volviendo al enfrentamiento entre el funcionamiento de la visión humana y artificial, se establece la comparación de ambas a continuación:

- Ojo humano

Por una parte el ser humano recibe información visual tal y como se describiría en un mundo analógico, es decir, de forma continua.

Sin entrar en demasiado detalle en el ámbito anatómico, la visión en el hombre se explica como la capacidad del ojo para detectar la luz y transformar la energía lumínica en señales eléctricas las cuales viajan al cerebro mediante el nervio óptico. Entre sus componentes principales se encuentra el cristalino, una lente ajustable según la distancia al objetivo así como un "diafragma" denominado pupila, cuyo diámetro está regulado por el iris, y la retina que se trata del tejido sensible

a la luz. El funcionamiento del ojo se explica porque la luz atraviesa la pupila y el cristalino y se proyecta sobre la retina, zona en la que unas células fotorreceptoras la transforman en impulsos nerviosos que se trasladan, a través del nervio óptico, al cerebro.

Se ve claramente un tipo de señal analógica presente en el proceso, la luz, o mejor dicho la intensidad de la misma que puede representarse por la luminancia, es decir, candela por metro cuadrado.

Sensor artificial

No se puede aplicar de forma directa el concepto del funcionamiento del ojo humano en un sensor artificial ya que trabaja con información digitalizada.

De este modo hace falta unos pasos intermedios antes de llegar a un resultado final. Por lo tanto hay que utilizar una aproximación diferente para resolver el problema y es aquí donde entra el concepto de nube de puntos.

1.2. Concepto de nube de puntos

Una nube de puntos se puede definir como una estructura P que representa un conjunto de puntos multidimensionales $p \in \mathbb{R}^n$. En el caso de una nube de puntos en tres dimensiones, cada elemento o punto está representado como mínimo por sus coordenadas geométricas X, Y, Z respecto a un sistema de referencia dado. Pero se puede añadir más datos todavía en forma de color, curvatura o información sobre la normal n a una superficie en un ámbito local de la misma. *añadir imágenes de nubes de puntos, ejemplos visuales*

Por lo tanto una nube de puntos es un conjunto de puntos individuales sin relación alguna entre ellos, cuya posición, color y otro tipo de características tienen definición, edición y representación muy simple por lo que es realmente práctico y sencillo manejar una gran cantidad de ellos sin tener que preocuparse por conceptos como escala, rotaciones y demás relaciones entre diferentes puntos de un mismo objeto, solamente posición y color importan.

La información integrable en un anube de puntos es: espacial, color, reflectividad, densidad

Se puede derivar de este concepto una gran potencia y flexibilidad ya que si se tienen una cantidad suficiente de puntos dispuestos correctamente se pueden representar todo tipo de superficies aunque en realidad no se trate de un plano continuo, es decir, el cerebro es capaz de interpretar complejas formas a partir de un tipo de información tan sencilla como las tres coordenadas espaciales.

Es más, se pueden llevar a cabo conversiones para relacionar el conjunto de puntos de la nube y crear superficies reales en tres dimensiones. Este tipo de conversión se denomina también como reconstrucción de superficies, es

decir, se parte de información puntual y se crea una superficie continua estimando qué relación puede haber entre puntos cercanos. De esta forma, una nube de puntos puede transformarse a una malla de polígonos o triángulos o incluso modelos CAD.

Las técnicas de reconstrucción de superficies son variadas y entre ellas se encuentran la triangulación de delaunay, que construye una red de triángulos sobre los vértices de la nube de puntos, y el método de marching cubes, que trata de convertir la nube de puntos en un campo de distancias volumétricas para poder así reconstruir la superficie.

Por otra parte, una peculiaridad o limitación respecto a las nubes de puntos tiene que ver con que la información que representan es superficial, es decir, los puntos siempre pertenecen a la superficie del objeto en cuestión ya que es el lugar donde la luz de los escáneres se llega, rebota y devuelve la información correspondiente.

Otra desventaja inherente a las nubes de puntos es la interpretación de la información que contienen ya que se ha explicado que está compuesta de un conjunto de objetos o puntos sin relación entre ellos. Es aquí donde se requiere intervención del ser humano pues es su cerebro el que puede encontrar la similitud entre una nube de puntos dada y el objeto o escenario que se supone que representa. Existe software capaz de encontrar patrones y características para clasificar nubes de puntos pero nunca de forma completamente fiable (ver aliasing)

El concepto de nube de puntos es eminentemente simple así como versátil y de gran utilidad. Esto se puede apreciar con gran multitud de aplicaciones del concepto de nube de puntos en el mundo real y es que este progreso tecnológico es un gran paso adelante para refinar procesos ya existentes, desde producción a nivel industrial hasta establecer las bases de la navegación de cualquier robot o vehículo.

Ejemplos hacer nube de puntos simple con visor y poner captura comparando con cosas como ... (tamaño, número de puntos, color/no color...)

Una aplicación de la reconstrucción de superficies es la creación de modelos digitales de elevación dentro

de la disciplina de sistemas de información geográfica. En el ámbito de la medicina se utilizan las nubes de puntos para representar información volumétrica para observar el interior del organismo sin necesidad de radiografiar o intervenir quirúrgicamente... ...aplicaciones

- ROS: For robots to work in unstructured environments, they need to be able to perceive the world.

medicina

- digitalizar terreno (<http://www.mdpi.com/1424-8220/8/8/4505/htm>)
- reconstrucción entornos 3D y reconocimiento de objetos

- navegación
 - industria 4.0
- explicar brevemente cada uno*

1.3. Adquisición de información: Sensores y evolución

La adquisición de información se lleva a cabo mediante sensores ópticos que lanzan rayos de luz (visible o infrarroja en la mayoría de los casos) que desvelan la información sobre el objeto escaneado al rebotar contra el mismo.

El factor clave es entonces el hecho de que estos rayos de luz puedan llegar a toda la superficie del objeto que se quiere analizar, es decir, accesibilidad física del sensor. De este modo, independientemente del sensor o método que se utilice, es imposible recolectar información sobre superficies no visibles o lo que es lo mismo, con un solo escaneo. Como consecuencia, es necesario llevar a cabo varios escaneos desde diferentes puntos de vista y ponerlos en conjunto conociendo con precisión la posición del sensor en cada escaneo.

En los últimos veinte años, se han hecho grandes progresos ya que actualmente se usan de sofisticadas cámaras y escaners láser y se han dejado atrás los sensores basados en sonar o infrarrojos los cuales proporcionan a penas unos bytes de información sobre su entorno.

Para entender la potencia de estos avanzados sensores se comparan tres tipos de tecnologías:

definición-descripción-aplicaciones/importancia

- Lidar

Definición: (light detection and ranging pero originalmente se conocía por la unión de light and radar) es un procedimiento que se originó a principios de la década de los sesenta tras la invención del láser y permite medir distancias a un objetivo iluminándolo con pulsos láser y estudiando los tiempos de retorno y longitudes de onda de estos pulsos al rebotar y ser captados por un sensor.

Este método combina precisión y versatilidad ya que puede valerse de luz visible, infrarroja o ultravioleta para lanzarla contra objetivos de diversos tipos de materiales como metal, cerámica, aerosoles, terreno (rocas y tierra) e incluso se puede llegar al nivel molecular

Descripción:

Existen dos formas de aplicar el método lidar:

incoherente:

- coherente:

En cuanto a los componentes necesarios para cualquier sistema lidar

Aplicaciones/dispositivos:

mencionar y describir alguna aplicación relevante ejemplo: Velodyne spinning LIDAR <https://en.wikipedia.org/wiki/Lidar>

Sonar

- Infrarrojo

Disponer de complejas, fiables y robustas representaciones del mundo real tiene no es tan sencillo como pueda parecer puesto que estos sensores suelen tener un precio prohibitivo para la mayoría de los interesados ya sean particulares o incluso empresas con un poder adquisitivo considerable. Sin embargo, las tornas han cambiado desde que han aparecido en el mercado ciertos sensores 3D como por ejemplo el sensor Kinect de la consola Xbox360 de Microsoft. Este sensor está basado en la tecnología PrimeSense y aunque puede trabajar con nubes de puntos en tiempo real e imágenes en 2D su precio no supera los 150\$. De este modo se ha producido un gran paso adelante en cuanto a los impedimentos relacionados con la adquisición, mantenimiento y delicadeza del hardware que traduce el mundo real a nubes de puntos.

1.4. Procesamiento software de información

Una vez estudiado el hardware necesario se precisa ahora de un mecanismo para trabajar con la inmensa cantidad de información que aportan los sensores.

El software existente para dicha tarea es diverso y no siempre gratuito. Como ejemplos se tienen:

- 3DF Zephyr

RealityCapture

- Agisoft Photoscan

- Point cloud tool

describir brevemente y mencionar impedimentos de precio, limitaciones etc

Es aquí donde entra en juego una librería de código abierto llamada PCL (insertar imagen-web) siglas que en inglés representan Point Cloud Library, traducible como librería de nubes de puntos.

Definir qué es pcl, historia, extensión <http://pointclouds.org/> paper: 3D is here

1.5. Descripción de herramientas

1.5.1. Herramientas hardware

1.5.2. Herramientas software

1.6. objetivos

Generar un programa capaz de obtener keypoints a partir de una nube de puntos
Croscompile usando las herramientas de xilinx
Enviar ejecutable a la placa y ejecutar el programa (optimización)

Bibliografía

- [1] Hassan Khalil. (2). (1996). Nonlinear systems. New Jersey, USA: Prentice Hall.