

Universidad Autónoma Metropolitana
Unidad Azcapotzalco
División de Ciencias Básicas e Ingeniería
Ingeniería en Computación

Propuesta de un modelo de adquisición, almacenamiento y representación
de Imágenes en 3D estereoscópico

Hugo Romero Romero - 209386625

Trimestre 10-I

29 de Marzo de 2010

Me comprometo a asegurar que el alumno cuente con los recursos materiales necesarios para que esta propuesta se realice satisfactoriamente.



Dr. Carlos Avilés Cruz



Dr. Juan Villegas Cortez

Objetivos Generales

Analizar, diseñar e implementar un método de adquisición y representación de imágenes estereoscópicas hacia una representación de imagen en 3D.

Objetivos particulares

Desarrollar un formato de almacenamiento de video estereoscópico con la capacidad de almacenar y comprimir dos flujos de video.

Generar un algoritmo de codificación y decodificación para el formato de video propuesto, el cual pueda reconstruir y procesar los dos flujos de video hacia un único flujo, brindando la percepción de video en 3D.

Analizar y reportar por escrito el desempeño del algoritmo propuesto.

1 Antecedentes

La visión binocular humana es un proceso complejo en el cuál se unen las imágenes de dos dimensiones que se proyectan sobre la retina del ojo. Estas imágenes presentan pequeñas diferencias entre sí, debido a la separación entre los ojos (Figura 1), que varía en promedio 65 mm. Esto es lo que proporciona el sentido visual de volumen y profundidad en los objetos y realidad percibidos. Este proceso de reconstrucción de percepción de profundidad y volumen se le conoce como percepción tridimensional o 3D [2].

Hablando un poco de historia, la palabra *estereoscopia* proviene del griego *stereos* y *skopen* que significa “observación sólida”, y hace referencia a la capacidad de observar objetos tridimensionales con cierto nivel de profundidad [7].

La idea de la estereoscopia se remonta al año 1838, en un documento de Charles Wheatstone, llamado “Algunos fenómenos trascendentes de la visión binocular” donde menciona las diferencias en los objetos percibidos dependiendo del ángulo de los ejes ópticos y cómo un par de imágenes pueden ser representadas por el cerebro como una sola[9].

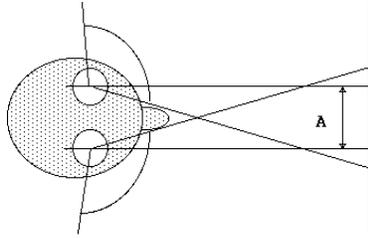


Figura 1. Visión humana (A) Paralaje de la visión binocular

Métodos para la visualización estereoscópica.

Los sistemas de visión estereoscópica se dividen en métodos pasivos y activos. Los métodos pasivos requieren de algún dispositivo especial para su visualización, los cuales suelen ser lentes que permiten dirigir cada imagen al ojo correspondiente. Algunos de estos sistemas pasivos pueden ser:

- sistema anáglifo, en el que las imágenes de dos dimensiones generan dos vistas en perspectiva separadas en un par de colores complementarios [8],
- el sistema polarizado, que puede ser lineal o circular, donde se observa a través de un par de filtros con la polarización correcta para cada ojo,
- el sistema temporal donde se muestran de forma alternada las imágenes correspondientes para cada ojo, donde los lentes son obturadores que alternan su estado permitiendo la visión de un solo ojo [10], entre otros.

Los sistemas activos permiten la visualización de imágenes estereoscópicas sin necesidad de visores especiales, dentro de los que se encuentran las pantallas auto-estereoscópicas .

2 Justificación

La compresión de video estereoscópico es importante, debido a que el ancho de banda necesario para el almacenamiento y transmisión incrementa linealmente con el número de canales de video implementados.

Los estándares de video MPEG¹ soportan el almacenamiento y transmisión de múltiples flujos de video, los cuales pueden ser aprovechados para la obtención de una imagen en 3D estereoscópico.

¹ Formato de video propuesto por el grupo de expertos en imágenes móviles (Moving Picture Experts Group)

Actualmente la tendencia en video se presenta en el desarrollo de sistemas de proyección en cine y televisión en 3D. Existen en el mercado diferentes propuestas por parte de diversos fabricantes del ramo de cine y video digital que ahora buscan incursionar en este mercado con productos que brinden al espectador la sensación de una visión en tres dimensiones (Panasonic 3D FullHD camera, Sony Alpha 3D SLR camera, Fuji FinePix REAL 3D W1, etc.); pero no existe un estándar para la captura, codificación, almacenamiento y reproducción de este tipo de material (video 3D). Cada una de estas propuestas son experimentales aún y ninguna de ellas ha demostrado ser la mejor respecto a las demás.

Considerando lo anteriormente expuesto se plantea la posibilidad de incursionar en una propuesta de desarrollo para la captura de video estereoscópica, su almacenamiento, codificación y decodificación; como una alternativa realizada a partir del estudio de las bases del procesamiento digital de imágenes y el reconocimiento de patrones.

3 Descripción Técnica

El proceso de visión humana es de tipo estereoscópica, la forma en que se simula por medio de los sistemas de adquisición de video o imágenes es a partir de dos dispositivos de captura alineados y con una separación que simule la distancia de separación de los ojos. Es gracias a esta separación que las imágenes adquiridas en paralelo tienen una diferencia sutil que permite obtener las indicaciones de profundidad.

A continuación se muestra, en la figura 2, el diagrama a bloques de las fases en las que se dividirá el proyecto. Las etapas mostradas en el diagrama se detallan líneas abajo.

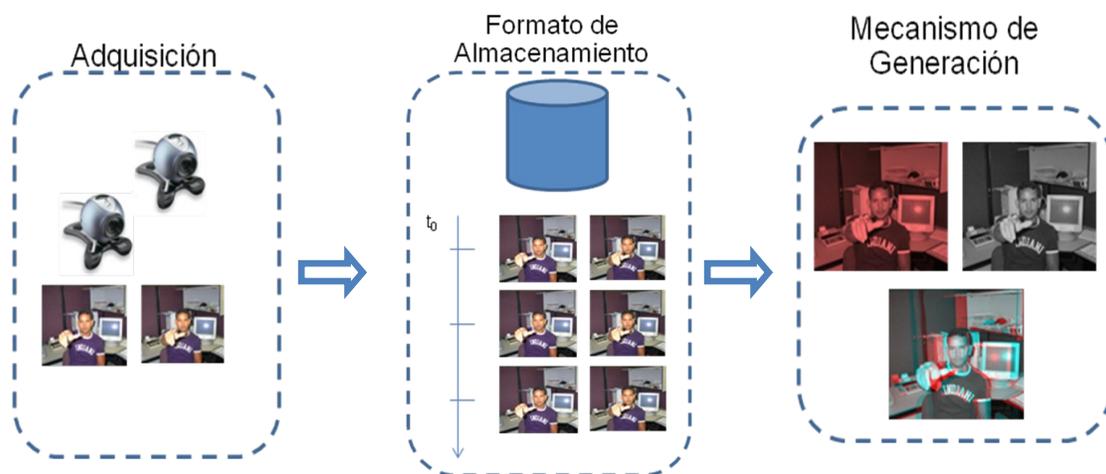


Figura 2. Diagrama de metodología tentativa para la propuesta.

Los productos a obtener en este proyecto se encuentran concentrados hacia el bloque de formato de almacenamiento debido a que son productos de software, sin embargo para poder comprobar el correcto funcionamiento de éstos, es necesario participar en la adquisición así como en la generación de estas imágenes.

Para la generación de señales visuales estereoscópicas se deben seguir los siguientes pasos: captación del par de imágenes, compresión de las señales digitales y despliegue de las imágenes.

El proceso de adquisición se llevará a cabo por medio de dos cámaras web, las cuales necesitan guardar una alineación dentro de los tres ejes (horizontal, vertical y profundidad) controlando el enfoque y el punto de convergencia de los ejes ópticos, cuyo flujo de video será adquirido en el equipo de cómputo que realizará el procesamiento para el almacenamiento del archivo. Al momento de la captura se aplicará el algoritmo de codificación de tal manera que el archivo quede en el formato especificado, quedando listo para la reproducción.

El formato de almacenamiento (el algoritmo a desarrollar) debe ser capaz de comprimir las señales digitales, determinando la información que puede ser eliminada a partir de la redundancia estadística. La redundancia estadística, espacial o temporal se calcula a través de la repetición de píxeles (semejantes en sus tonos de color) dentro de una imagen y entre las imágenes secuenciadas. El formato de almacenamiento debe permitir la identificación de la correspondencia en tiempo entre cada par de imágenes capturado.

El mecanismo de generación, i.e. el despliegue de las imágenes estereoscópicas, tiene como objetivo principal el brindar la imagen correspondiente para formar la percepción de tridimensionalidad.

Los métodos activos requieren de visores especiales con componentes electrónicos para poder visualizar de forma correcta el flujo de imágenes decodificado. Los pasivos no requieren de componentes electrónicos externos para una correcta visualización por lo que se considera trabajar con la opción de anáglifo por filtrado de color debido al bajo coste de implementación ya que sólo requiere de unos lentes de color.

El método anáglifo consiste en la extracción de las bandas de color dentro del formato RGB² para cada fuente de la imagen y la síntesis en la imagen final. El estándar más utilizado es el rojo-cyan el cuál es un par de colores

²Formato de tratamiento de color en bandas de color Rojo, Verde, Azul (Red, Green, Blue).

complementarios en el círculo cromático, que consiste en aislar la imagen de origen izquierda en su banda de color rojo, posteriormente se elimina la banda de color rojo de la imagen de origen derecha permaneciendo las bandas de color verde y azul. Finalmente se asignan las bandas resultantes en una sola imagen.

Se pretende trabajar en un ambiente Windows, con un lenguaje de programación capaz de interactuar con los controladores de recepción del flujo de datos de las cámaras, utilizando la biblioteca en openCV para el ajuste de captura y el procesamiento.

4 Especificaciones Técnicas

El desarrollo del proyecto se dirige al formato de almacenamiento independientemente del método de adquisición, suponiendo solamente que se tienen dos fuentes de origen.

Las imágenes y video serán obtenidos de forma y condiciones controladas en el laboratorio de multimedia para controlar las variaciones producidas por los cambios de iluminación.

Se utilizará el método análogo para comprobar el correcto funcionamiento del algoritmo, y se estudiará la posibilidad de extenderse a otros métodos de presentación.

Se dará por concluido el proyecto al obtener la decodificación adecuada de diferentes pruebas de video, decodificadas usando el algoritmo desarrollado.

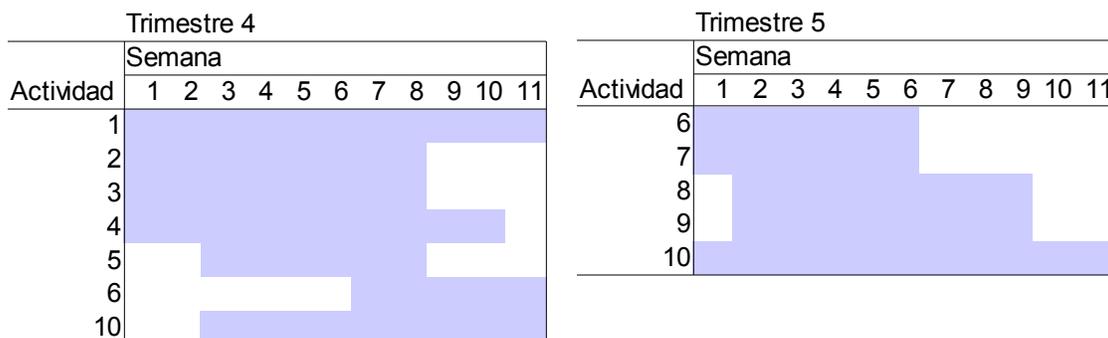
5 Recursos

Actualmente se cuenta con el equipo de cómputo para el procesamiento de imágenes en el Laboratorio de Multimedia:

- Equipos Workstation con procesador Quad-core Intel Xeon 5500, de 64 bits, con tarjetas de video Nvidia.
- Cámaras web de video.
- Condiciones apropiadas para el desarrollo de los experimentos.

6 Plan de trabajo

1. Estudio de las técnicas de representación de imágenes estereoscópicas.
2. Estudio de métodos de compresión de imagen y video (JPEG, MPEG).
3. Estudio de reconocimiento de patrones.
4. Estudio de los métodos estandarizados de adquisición y representación de video (MPEG).
5. Preparación del sistema de captura de video y realización de capturas de muestras en condiciones controladas.
6. Codificación del software de captura y almacenamiento.
7. Codificación de software de decodificación de video.
8. Pruebas de aplicación a las muestras de video.
9. Análisis de resultados y comparación de desempeño por dimensionalidad.
10. Redacción de la tesis.



7 Bibliografía

1. Acosta, G. y Ríos, H. V. (1996). "Módulo de visión 3D a partir de reconstrucción estereoscópica". *Memorias del IV Simposio Nacional de Sistemas Computacionales*, Instituto Tecnológico de San Luis Potosí.
2. Blanchard, J., and Tsuneto, R. 1993. Stereoscopic viewing. *Encyclopedia of Virtual Environments*. World Wide Web URL: <http://www.hitl.washington.edu/scivw/EVE/III.A.1.b.StereoscopicViewing.html> .
3. Gonzalez, R. C. and Woods, R. E. (2008). "Digital Image Processing", 3rd ed., Prentice Hall, Upper Saddle River, NJ.
4. Ohm, J.R. (1999). "Stereo/multiview video encoding using the MPEG family of standards". In: Merritt, O.J., Bolas, M.T., Fisher, S.S., (eds.) *The Engineering Reality of Virtual Reality*, vol. 3639, pp. 242–253. SPIE, San Jose.
5. Ohm, J.R. Karsten Müller. (1998). "Incomplete 3D for multiview representation and synthesis of video objects". *Multimedia Applications, Services and Techniques. Lecture Notes in Computer Science Vol. 1425*.
6. T. Sikora. (1997). "MPEG digital video coding standards", *IEEE Signal Proc. Mag.*, vol. 14, no. 5, pp. 82-100.
7. Valverde V. F., Peñaherrera H. W., Pérez R. T. (2005). Estudio del sistema de televisión estereoscópica como una aplicación de la televisión digital. *XIX jornadas en ingeniería eléctrica y electrónica Vol. No. 19*
8. Volbracht, S. ; Shahrabaki, K. ; Domik, G. ; Fels, G. Perspective viewing, anaglyph stereo or shutter glass stereo?
9. Wheatstone C. *Contributions to the Physiology of Vision.—Part the First. On some remarkable, and hitherto unobserved, Phenomena of Binocular Vision* "Philosophical Transactions" of the Royal Society of London, Vol. 128, pp. 371 - 394.
10. Binocularity: The staff and students of the Durham Visualization Laboratory (DVL): <http://www.binocularity.org/stereoscopic.php>