****

**FACULTAD DE INGENIERÍA**

**Vicedecanatura Académica**

**POSGRADOS**

**PRESENTACIÓN PROPUESTA**

TESIS DE MAESTRÍA: TRABAJO FINAL DE ESPECIALIZACIÓN:

1. PROPONENTE: Alexander Pinzón Fernández CÓDIGO: 02 299802
2. PROGRAMA: Maestría en Ingeniería de Sistemas y Computación
3. DIRECTOR PROPUESTO: Eduardo Romero Castro

DEPARTAMENTO: Ingeniería de Sistemas e Industrial.

1. TÍTULO: Extracción y seguimiento del esqueleto de un cuerpo en movimiento a partir de múltiples vistas.

ÁREA: Ciencias de la Computación

LÍNEA DE INVESTIGACIÓN: Computación grafica, procesamiento de imágenes, procesamiento de video.

1. ANTECEDENTES Y JUSTIFICACIÓN:

El estudio de métodos para la captura de movimiento del cuerpo humano y sus partes es de gran interés para la animación computarizada, los ambientes virtuales interactivos, la manufactura de dispositivos para la interacción hombre maquina, el análisis cinemático para el diagnostico de patologías asociadas con la marcha, entre otros (1).

La extracción y seguimiento de esqueletos se logra mediante la utilización de un conjunto de técnicas organizadas dentro de un proceso. Tales técnicas son la adquisición de datos desde un objeto real, la reconstrucción 3D a partir de los datos de la captura, la extracción del esqueleto del objeto tridimensional reconstruido, y el seguimiento del esqueleto tomado como un proceso iterativo de captura con un esqueleto invariante topológicamente en el tiempo.

Los esqueletos son estructuras 1D que representan una simplificada versión de la geometría y topología de un objeto 3D. Son en esencia una entidad geométrica que abstrae el volumen del objeto. Por esta razón muchos métodos de extracción del esqueleto requieren una representación discreta del volumen del modelo de entrada. Sin embargo la mayoría de modelos usados en aplicaciones de computación grafica son disponibles únicamente como una representación de su superficie, tales como las mallas de polígonos (2).

5.1. Adquisición de datos

Existen diferentes enfoques en la adquisición de la información espacial de objetos reales, algunos basados en el procesamiento de imágenes, otros basados en sensores (mecánicos, electromagnéticos, e inerciales entre otros), y otros como el escáner de laser, la tomografía, el ultrasonido, etc (3).

Los métodos de adquisición basados en el procesamiento de señales utilizan las secuencias de imágenes producidas por una o varias cámaras. Cuando se utilizan varias cámaras se explota la coherencia espacial dada por la visión estereoscópica, mediante el uso de puntos característicos como marcadores (luminosos, reflectivos, o detectados mediante filtros), y las siluetas de los objetos obtenidos mediante la segmentación de las imágenes (4). Cuando se utiliza una única imagen se hace uso de modelos a priori y además se añaden restricciones en las búsquedas de soluciones, estos modelos pueden basarse en aproximaciones probabilísticas o aprendizaje de máquina.

5.1.2. Segmentación de imágenes

La segmentación de imágenes son un conjunto de métodos para dividir una imagen en regiones, dadas ciertas características. La segmentación puede ser automática es decir sin interacción del usuario, o semiautomáticas que requieren la supervisión manual (5). A continuación mostraremos algunos de los métodos de segmentación más usados:

* Métodos de segmentación de contorno activos (6).
* Método Snake con uso de filtros de gradiente de dirección (7).
* Métodos que utilizan la coherencia espacio temporal de formas limitantes especialmente cajas y elipses (8).
* Métodos probabilísticos con redes bayesianas para modelar la interacción de campos (9).

5.1.3. Captura mediante sensores

Este tipo de técnicas ofrecen gran precisión y libertad para la captura del movimiento, pues no presentan problemas de oclusión presentes en los métodos de procesamiento de imágenes (10).

Algunas desventajas son:

* No proporcionan información de la forma y superficie del objeto.
* Es necesario el uso de trajes especiales que restringen la captura de la textura del cuerpo. Además algunos mecanismos no son inalámbricos lo que restringe la movilidad.
* Son equipos costosos.
* Los sistemas requieren estar próximos a los objetos, no permitiendo grandes translaciones en el espacio (11).

5.1.4 Escáner tridimensional de laser

Ofrecen la mayor calidad y precisión en la captura de la forma y detalles de la superficie del objeto, siendo de los métodos más utilizados en industrias como la del cine, la paleontología, la arqueología, entre otros (12). La mayor desventaja de esto método son los altísimos costos de estos dispositivos.

5.2. Reconstrucción tridimensional

La reconstrucción tridimensional, son un conjunto de técnicas mediante las cuales objetos reales, son representados mediante estructuras de datos o funciones, haciendo énfasis en características como volumen y forma. Existe actualmente una gran cantidad de técnicas de reconstrucción y métodos de mallado 3D (13), dado un conjunto de puntos 3D no conectados.

5.2.1 Reconstrucción basada en imágenes

La reconstrucción basada en imágenes, tiene dos enfoques importantes uno con el uso de marcas, y otro sin ellas, además hay que tomar en cuenta los parámetros intrínsecos (distancia focal, distorsión, etc.) y extrínsecos (posición, orientación, etc.) de la cámara (14).

En la reconstrucción con el uso de marcas se hace un filtrado de la imagen para localizar estos marcadores los cuales son en general puntos de algún color poco presente en la escena como el blanco, un rojo o verde intenso, también se usan marcadores que emiten luz como pequeños bombillos, o el uso de imágenes en blanco y negro que tienen un patrón dado reconocible. Con estos puntos segmentados y localizados en las imágenes se procede a realizar la solución de las ecuaciones de las transformaciones lineales correspondientes basadas en un sistema estéreo con un modelo de cámara pinhole (15).

5.3 Extracción del esqueleto

La extracción del esqueleto es un problema que actualmente es tratado desde distintos enfoques, algunos realizan estimaciones del esqueleto en 2D, y luego utilizan información a priori para estimar como fue transformado el esqueleto tridimensional real (16), otros enfoques realizan una reducción iterativa del número de polígonos con ciertas restricciones de contracción y minimización de energía, hasta encontrar un esqueleto irreducible (2). Otros métodos utilizan distintas representaciones topológicas de un objeto tridimensional mediante grafos, luego realizan una simplificación de este grafo, y por ultimo utilizan un esqueleto modelo el cual transforman para minimizar el error entre el esqueleto y el grafo dado por medio de alguna métrica basada en heurísticas (12).

1. IDENTIFICACIÓN DEL PROBLEMA:

La extracción del esqueleto de un cuerpo articulado basado en imágenes, se realiza tradicionalmente por medio del seguimiento de marcas desde una o varias cámaras.

En este trabajo se busca extraer y capturar el movimiento de un esqueleto reconstruido a partir de una nube de puntos tridimensionales estimados desde las imágenes de varias cámaras.

1. OBJETIVO GENERAL Y OBJETIVOS ESPECÍFICOS:

**Objetivo General:**

Extracción y captura de movimiento del esqueleto de un cuerpo articulado, observado desde varias cámaras al mismo tiempo.

**Objetivos Específicos:**

* Adquirir las secuencias de imágenes del cuerpo articulado observado.
* Segmentar las siluetas y puntos característicos desde las secuencias de imágenes.
* Realizar una estimación tridimensional de de los puntos característicos.
* Realizar una estimación del volumen del cuerpo a partir de las siluetas.
* Extraer el esqueleto a partir de la información tridimensional estimada.
* Realizar el seguimiento del esqueleto en cada cuadro de los videos.
* Comparar los resultados obtenidos de la extracción y seguimiento del esqueleto.
* Implementación de un sistema de realidad aumentada con una interfaz de usuario que permita visualizar el esqueleto en movimiento, junto con el video de cualquiera de las cámaras de la escena.

1. METODOLOGÍA:
   1. Exploración de las investigaciones y avances realizados en el área.
   2. Adquisición de Datos: Obtener la secuencia de imágenes utilizando 6 cámaras con las siguientes descripciones:
      * Cámaras geométricamente localizadas en un arreglo circular alrededor del centro de la escena.
      * Cámaras con características intrínsecas y extrínsecas calibradas.
      * Cámaras sincronizadas en el tiempo que capturan la escena desde distintos ángulos.
      * Escenario con condiciones de iluminación controladas.
   3. Extracción de la silueta del cuerpo articulado: seleccionar e implementar una técnica de segmentación de video, adecuada para extraer la silueta del cuerpo articulado.
   4. Extracción de puntos característicos: seleccionar e implementar una técnica de segmentación de video adecuada para extraer puntos característicos de la geometría del cuerpo articulado.
   5. Estimación tridimensional de los puntos característicos: selección e implementación de una técnica para la estimación tridimensional de los puntos característicos, este método debe tomar en cuenta la información extraída de las otras cámaras y sus parámetros de calibración, para determinar la correspondencia y localización tridimensional de los puntos característicos en cada una de las secuencias.
   6. Estimación tridimensional del volumen: seleccionar e implementar una técnica para la estimación del volumen del cuerpo articulado a partir de las siluetas extraídas desde las secuencias de video, junto con los parámetros de calibración de las cámaras.
   7. Extraer el esqueleto: Desarrollo e implementación de una técnica para extraer el esqueleto de un cuerpo articulado dados los puntos característicos, el volumen del objeto y los parámetros de las cámaras.
   8. Seguimiento del esqueleto: Desarrollo e implementación de una técnica para determinar la correspondencia y cambio de pose del esqueleto a través del tiempo en secuencias de video.
   9. Comparación de los resultados: se realizara comparando los resultados de la extracción del esqueleto con otros métodos de extracción del esqueleto, utilizando las mismas secuencias de video.
   10. Implementación de un sistema de realidad aumentada: Realizar los pasos necesarios para la construcción de un sistema basado en técnicas de ingeniería de software.
2. ACTIVIDADES A DESARROLLAR:

* Recopilar y analizar información para escribir el estado del arte.
* Adquisición de los datos:
  + Ubicar las cámaras en un arreglo circular, alrededor del centro de la escena.
  + Iluminar el escenario de acuerdo a parámetros seleccionados.
  + Calibrar las cámaras.
  + Capturar de forma sincronizada los videos de un cuerpo articulado en movimiento.
* Segmentación de la silueta del cuerpo articulado:
  + Selección de una técnica de segmentación de imágenes.
  + Implementar la técnica de segmentación.
* Detección de puntos característicos:
  + Selección de una técnica para la detección de puntos característicos.
  + Implementar la técnica de detección.
* Estimación tridimensional de los puntos característicos:
  + Seleccionar una técnica de reconstrucción tridimensional.
  + Implementar la técnica de reconstrucción tridimensional.
* Estimación tridimensional del volumen:
  + Seleccionar una técnica de reconstrucción tridimensional desde las siluetas.
  + Implementar la técnica de reconstrucción tridimensional desde las siluetas.
* Extraer el esqueleto:
  + Desarrollar la técnica de extracción del esqueleto.
  + Implementar la técnica de extracción del esqueleto.
* Seguimiento del esqueleto:
  + Desarrollar la técnica de seguimiento.
  + Implementar la técnica de seguimiento.
* Implementación de un sistema de realidad aumentada:
  + Realizar un análisis de requerimientos
  + Hacer un diseño del sistema.
  + Implementar el diseño.
  + Documentar el sistema.
* Comparar los resultados obtenidos
  + Seleccionar otros métodos de extracción del esqueleto que usen secuencias de video desde varias cámaras.
  + Obtener los resultados de los otros métodos.
  + Comparar los resultados del método desarrollado en esta tesis con los otros seleccionados.
* Documentación.
  + Escritura del documento de tesis.
  + Elaboración de artículos.

1. CRONOGRAMA: Vea la Figura 1.



Figura 1: Cronograma para el desarrollo de las actividades

1. BIBLIOGRAFÍA BÁSICA:

1. *3-D Reconstruction of Static Human Body Shape from Image Sequence.* **Remondino, Fabio.** 2004, Computer Vision and Image Understanding, Vol. 93, págs. 65-85.

2. *Skeleton Extraction by Mesh Contraction.* **Au, Oscar Kin-Chung, y otros.** 2008, ACM Transactions on Graphics, Vol. 27, pág. 10.

3. *Visual hull alignment and refinement across time: a 3D reconstruction algorithm combining shape-from-silhouette with stereo.* **Cheung, G.K.M., Baker, S. y Kanade, T.** 2003. Vol. 2, págs. II-375-82 vol.2.

4. *3D reconstruction by combining shape from silhouette with stereo.* **Lin, Huei-Yung y Wu, Jing-Ren.** 2008. págs. 1-4.

5. *Image-based shape model for view-invariant human motion recognition.* **Jin, Ning y Mokhtarian, F.** 2007. págs. 336-341.

6. *Human video textures.* **Flagg, Matthew, y otros.** s.l. : ACM, 2009. págs. 199-206.

7. *Skeleton Extraction of 3D Objects with Radial Basis Functions.* **Ma, Wan-Chun, Wu, Fu-Che y Ouhyoung, Ming.** s.l. : IEEE Computer Society, 2003. pág. 207.

8. *Shape from inconsistent silhouette for free viewpoint video.* **Landabaso, J.-L., Lizcano, L. y Pardas, M.** 2008. págs. 213-216.

9. *Methods for Volumetric Reconstruction of Visual Scenes.* **Slabaugh, Gregory G., y otros.** 2004, International Journal of Computer Vision, Vol. 57, págs. 179-199.

10. *Interaction capture and synthesis.* **Kry, Paul G. y Pai, Dinesh K.** s.l. : ACM Press, 2006, ACM Trans. Graph., Vol. 25, págs. 872-880.

11. *Multi-stereo 3D object reconstruction.* **Esteban, C. Hernandez y Schmitt, F.** 2002. págs. 159-166.

12. *Performance capture from sparse multi-view video.* **Aguiar, Edilson de, y otros.** s.l. : ACM, 2008. págs. 1-10.

13. *Reconstructing Human Shape and Motion from Multi-View Video.* **Aguiar, Edilson de, y otros.** s.l. : The IEE, 2005. págs. 42-49.

14. *Regression-based Hand Pose Estimation from Multiple Cameras.* **de, Teofilo E. y Murray, David W.** s.l. : IEEE Computer Society, 2006, Computer Vision and Pattern Recognition, IEEE Computer Society Conference on, Vol. 1, págs. 782-789.

15. *3D-modeling by ortho-image generation from image sequences.* **Thorm\"{a}hlen, Thorsten y Seidel, Hans-Peter.** s.l. : ACM, 2008. págs. 1-5.

16. **Remondino, Fabio y Roditakis, Andreas.** 3D Reconstruction of Human Skeleton from Single Images or Monocular Video Sequences. *3D Reconstruction of Human Skeleton from Single Images or Monocular Video Sequences.* 2003. págs. 100-107.

RECURSOS FÍSICOS:

|  |  |
| --- | --- |
| **Recurso** | **Dependencia** |
| Computador estación de trabajo | Estudiante |
| Conexión a internet | Universidad Nacional |
| Video Cámaras | Centro de Telemedicina |
| Software de procesamiento | Software Libre |
| Herramientas de desarrollo de software | Software Libre |
| Textos y artículos electrónicos | Universidad Nacional |
| Papelería y fotocopias | Estudiante |

Figura 2. Recursos Físicos.

Este proyecto se desarrollara con los recursos expuestos en la Figura 2., en las instalaciones del Centro de Telemedicina de la Universidad Nacional de Colombia.

1. COSTOS DEL TRABAJO Y FUENTES DE FINANCIACIÓN:

Ver el cuadro en la Figura 3.



Figura 3. Costos del trabajo y fuentes de financiación

1. COMENTARIO CON VISTO BUENO DEL DIRECTOR: (calificar los siguientes aspectos: organización, pertinencia, relevancia y originalidad).
2. FIRMA DEL PROPONENTE
3. FIRMA DEL DIRECTOR (ASESORES)
4. FECHA