

Resumen

La presente tesis trata sobre el estudio de técnicas de visión para la detección de la postura del esqueleto del cuerpo humano basada en el análisis de una sola imagen, además del seguimiento de estas posturas a lo largo de una secuencia de imágenes.

Se propone modelar la postura del esqueleto cuerpo humano mediante cuatro cadenas cinemáticas que modelan las cuatro extremidades articuladas. Estas cadenas cinemáticas y la cabeza permanecen unidas al cuerpo. Las cuatro cadenas cinemáticas se componen de tres puntos de interés. Por lo tanto, el modelo inicialmente dispone de un total de 14 puntos de interés.

Todos los métodos propuestos en este trabajo son implementados, validados y analizados utilizando la base de datos CAD60. La base de datos CAD60 es de dominio público y contiene imágenes en RGB y profundidad, junto con una base de datos creada expresamente para complementar a la base de datos CAD60.

En esta tesis se propone modificar la técnica denominada Deformable Parts Model (DPM), añadiendo el canal de profundidad denominado “Depth”. Inicialmente el modelo DPM se definió sobre imágenes de tres canales RGB. Mientras que en esta tesis se propone trabajar sobre imágenes de cuatro canales RGBD, por ello a la ampliación propuesta se le denomina 4D-DPM. Los experimentos realizados con 4D-DPM demuestran una mejora en la precisión de la detección de la postura con respecto al modelo inicial DPM, a costa de incrementar su coste computacional al tratar un canal adicional.

Por otra parte, se propone reducir el coste computacional anterior simplificando el modelo que define la postura del cuerpo humano. La idea es reducir el número de variables a detectar con el modelo 4D-DPM, de tal manera que las variables suprimidas se puedan calcular a partir de las variables detectadas, utilizando modelos de cinemática inversa basados en cuaterniones duales. Los experimentos realizados demuestran que la combinación de estas dos técnicas permite, reduciendo el coste computacional del método original DPM, mejorar la precisión de la detección de postura debido a la información extra del canal de profundidad.

Adicionalmente, se propone utilizar modelos de filtros de partículas para continuar mejorando la precisión de la detección de las posturas humanas a lo largo de una secuencia de imágenes.

Atendiendo al problema de detección y seguimiento de la postura del esqueleto del cuerpo humano a lo largo de una secuencia de vídeo, esta tesis propone el uso del siguiente método.

1. Calibración de cámaras. Procesamiento de imágenes RGBD. Sustracción del fondo de la imagen con el método MSER.
2. 4D-DPM: método utilizado para detectar los puntos de interés (variables del modelo de postura) dentro de una imagen.
3. Filtros de partículas: se diseña este tipo de filtros para realizar el seguimiento de los puntos de interés a lo largo del tiempo y corregir los datos obtenidos por el sensor.
4. Modelado cinemático inverso: se realiza el control de cadenas cinemáticas con la ayuda de cuaterniones duales con el fin de obtener el modelo completo de la postura del esqueleto del cuerpo humano.

La contribución global de esta tesis es la propuesta del método anterior que, combinando los métodos anteriores, es capaz de mejorar la precisión en la detección y el seguimiento de la postura del esqueleto del cuerpo humano en una secuencia de vídeo, reduciendo además su coste computacional.

Esto es posible debido a la combinación del método 4D-DPM con la utilización de técnicas de cinemática inversa. El método original DPM debe detectar 14 puntos de interés sobre una imagen RGB para estimar la postura de un cuerpo humano. Sin embargo, el método propuesto, donde se elimina un punto de interés por cada extremidad, debe detectar 10 puntos de interés sobre una imagen RGBD. Posteriormente, los 4 puntos de interés eliminados se calculan

mediante la utilización de métodos de cinemática inversa a partir de los 10 puntos de interés calculados.

Para resolver el problema de la cinemática inversa se propone utilizar cuaterniones duales para cada una de las 4 cadenas cinemáticas que modelan las extremidades del esqueleto del cuerpo humano.

El filtro de partículas se aplica sobre la secuencia temporal de los 10 puntos de interés del modelo de postura detectados a través del método 4D-DPM. Para diseñar estos filtros de partículas se propone añadir las siguientes restricciones para ponderar las partículas generadas:

1. Restricciones en los límites de articulaciones: la postura del esqueleto del cuerpo humano se modela con un conjunto de cadenas cinemáticas abiertas. De tal manera que los puntos de interés son las variables de articulación de cada una de las cadenas cinemáticas. Cada una de estas variables tiene un movimiento restringido en un rango determinado.
2. Restricciones de suavidad: se propone ponderar las partículas de manera inversamente proporcional a la distancia de la partícula generada con la solución en el instante de tiempo anterior.
3. Detección de colisiones: el modelado geométrico utilizado para modelar el esqueleto del cuerpo humano es un conjunto de poli-esferas ya que estas nos permiten realizar la detección de colisiones, entre los elementos del cuerpo, de manera muy eficiente. Se propone que el filtro de partículas no genere partículas en las cuales se produzcan colisiones imposibles entre estos elementos.
4. Proyección de las poli-esferas: se propone ponderar cada partícula generada de forma directamente proporcional al solapamiento de la proyección del modelo de las poli-esferas que define esta partícula con algún plano de la imagen RGBD.