Congr. Int. en Ing. Electrónica. Mem. ELECTRO, Vol. 39, pp. 293-299, Oct 2017, Chihuahua, Chih. México
http://electro.itchihuahua.edu.mx/memorias_electro/MemoriaElectro2017.zip
ISSN 1405-2172

MODELADO TRIDIMENSIONAL DE LA FORMA DE ANDAR USANDO DATOS FUSIONADOS DE DOS SENSORES RGBD

Isidro Robledo Vega, Víctor Hugo Velasco Fernández, Alberto Pacheco González

Tecnológico Nacional de México
Instituto Tecnológico de Chihuahua
División de Estudios de Posgrado e Investigación
Ave. Tecnológico #2909, Col. 10 de Mayo, Chihuahua, México
Tel. +52(614)201-2014
[irobledo, vhvelasco, apacheco]@itchihuahua.edu.mx

RESUMEN

En este artículo se describe un modelo de la forma de andar de personas construido a partir de secuencias de datos de profundidad capturados con dos sensores RGBD y fusionados en una sola nube de puntos. Se extrae un ciclo completo de la forma de andar de cada secuencia fusionada, las nubes de puntos correspondientes a este ciclo son acumuladas en un volumen de datos (VDA) para capturar los rasgos de la forma de andar de una persona. Se calcularon VDAs con las secuencias fusionadas de 20 personas y se utilizaron para entrenar y probar un sistema de reconocimiento basado en análisis de componentes principales logrando clasificar correctamente todos los VDA de prueba. Se realizó otro experimento similar pero utilizando los datos de un solo sensor RGBD para calcular los VDA, en este caso se logró clasificar correctamente 80% de los VDA de prueba. Con estos experimentos comprobamos que el modelo calculado con los datos de dos sensores RGBD captura una mayor cantidad de información del movimiento de las personas mejorando los resultados del reconocimiento de la forma de andar.

Palabras clave: Sensores RGBD, fusión de datos, análisis de la forma de andar.

ABSTRACT

In this paper, we describe a human gait model built from sequences of depth data captured with two RGBD sensors and fused into a single point cloud. One complete gait cycle was extracted from each fused sequence; the point clouds corresponding to this cycle are accumulated into a data volume (VDA) to capture the features of human gait. VDAs were computed with fused sequences from 20 people and used to train and test a PCA based recognition system managing to correctly classify all of the test VDAs. Another similar experiment was carried out but this time the data used to compute the VDA came from only one RGBD sensor. In this case only 80% of the test VDAs were correctly classified. With these experiments we verified that the model computed with data from two RGBD sensors captures a greater amount of human motion information improving results in human gait recognition.

Keywords: RGBD Sensors, data fusion, human gait analysis.

1. INTRODUCCIÓN

Una de las áreas de investigación más importantes dentro de Visión por Computadora es el análisis de movimiento, donde se investigan técnicas para la descripción y reconocimiento de objetos en movimiento, que pueden ser objetos rígidos o no-rígidos [1]. El modelo desarrollado y descrito en este artículo tiene que ver con el análisis de movimiento de objetos no-rígidos, en particular, el análisis de movimiento humano que se puede utilizar para la identificación de personas por su forma de andar.

Las tecnologías de identificación de personas son utilizadas en escenarios diseñados específicamente para verificación y control de acceso ofreciendo gran potencial para identificar individuos en escenarios complicados, tales como aeropuertos y otras instalaciones importantes. Muchos programas de investigación y desarrollo así como evaluaciones operacionales y técnicas buscan proporcionar elementos para toma de decisiones con una evaluación sobre la sostenibilidad de los sistemas de identificación en el estado del arte y su desempeño en diferentes entornos. El programa de identificación de personas a distancia, también llamado *HumanID Program* es uno de estos programas de desarrollo [2]. El programa *HumanID* participa en la investigación y desarrollo de una amplia variedad de tecnologías de identificación humana para mejorar el estado del arte. Sin embargo, resultados recientes de éste y otros programas sugieren que el desarrollo e implementación de un sistema robusto y fiable para identificación humana sigue siendo un gran reto. Existen tecnologías de identificación humana que han probado ser útiles en escenarios estructurados donde las características del ambiente y físicas son conocidas y controladas. En esta clase de escenario la tarea de identificación implica la recolección de información biométrica del sujeto y compararla con datos almacenados en una base de datos previamente recolectada. El resultado de esta comparación es utilizada en un proceso de decisión donde se permite o no acceso al sujeto en cuestión. Las tecnologías de identificación biométrica más utilizadas en escenarios controlados son iris, huellas dactilares, voz, geometría de la mano y rostro. Todas estas tecnologías requieren que el sujeto se encuentre a cierta proximidad o en contacto físico con el sensor biométrico para que este método sea efectivo.

El análisis de movimiento humano apareció hace unas décadas, principalmente con propósitos de análisis médico. Recientemente se han desarrollado variadas técnicas de análisis de la forma de andar, la mayoría de los trabajos están

enfocados en distinguir entre diferentes tipos de movimientos, tal como correr, caminar, trotar o subir escaleras, pero el uso de la forma de andar como característica biométrica para identificar personas ha recibido mayor interés por parte de los investigadores. Desde el punto de vista de vigilancia, el análisis de la forma de andar tiene gran interés ya que ésta puede obtenerse a distancia y sin llamar la atención [3].

Los métodos existentes para el reconocimiento de la forma de andar se pueden dividir en dos categorías, los métodos basados en modelos y los métodos basados en siluetas [4]. Estos últimos han demostrado ser más exitosos y de ellos el más sobresaliente de los últimos años ha sido el de Man y Bhanu [5] que caracteriza los patrones de la forma de andar utilizando el promedio de las imáges de las siluetas durante un ciclo completo de la forma de andar, refieriendose a la imagen promedio resultante como *Gait Energy Image* o GEI por sus siglas en inglés.

Los sistema de visión por computadora usados para la generación de datos tridimensionales del cuerpo completo de una persona en movimiento por lo general son muy costosos, lentos o demasiado complejos, como es el caso del escáner tridimensional y los sistemas de visión en estéreo. Gracias a la tecnología implementada en los dispositivos de captura de movimiento en complemento con los programas de software utilizados para algunos videojuegos, se han desarrollado sensores de bajo costo, llamados sensores RGBD (R-*Red* o rojo, G-*Green* o verde, B-*Blue* o azul y D-*Depth* o profundidad). Un sensor RGBD es un dispositivo que integra una imagen digital a color, en formato RGB, adquirida por una cámara de video digital a color con datos de profundidad obtenidos por medio de un sensor infrarrojo. Con su desarrollo se han generado grandes oportunidades para resolver problemas fundamentales en el área de visión por computadora, ya que permite el análisis y seguimiento de formas en 3D, y más específicamente de formas humanas [6].

El objetivo principal planteado en este trabajo fue la construcción de un modelo tridimensional de la forma de andar a partir de la fusión de datos obtenidos por medio del sistema de captura sincronizada de dos sensores RGBD descrito en [7]. Se utilizaron dos sensores con la finalidad de obtener una perspectiva del cuerpo humano desde dos puntos en el mismo instante de tiempo de manera que el modelo tridimensional sea más preciso ya que permite mantener una cantidad mayor de información sobre las características, posición y movimiento de una persona reduciendo las oclusiones.

2. ADQUISICIÓN Y FUSIÓN DE DATOS DE DOS SENSORES RGBD

2.1. Adquisición Sincronizada de Datos de dos sensores RGBD

Se llevó a cabo la adquisición de los datos de personas caminando utilizando un sistema de captura sincronizada con dos sensores RGBD. Los sensores son colocados uno frente a otro a 4 metros de distancia y la persona camina en líne recta en medio de los dos sensores, de esta forma se captura el movimiento de la persona con tomas frontales paralelas a cada sensor. El sistema cuenta con tres computadoras, dos de ellas denominadas Equipos de Captura y la otra denominada Equipo de Control. Los Equipos de Captura están conectados directamente a los sensores y una vez inicializados esperan la señal enviada desde el Equipo de Control para inicar la captura de datos.

Los sensores RGBD utilizados son sensores Kinect Versión 2 que proporcionan un flujo de datos de profundidad con una resolución de 512x424 puntos a una tasa de 30 cuadros por segundo. Se utilizaron funciones del SDK del sensor Kinect para mantener solo los datos de las persona en movimiento descartando el fondo. Los datos son almacenados con formato de nube de puntos transformados del espacio de coordenadas de la cámara de profundidad al espacio de coordenadas del mundo real utilizando como referencia el centro del sensor. Para cada punto se guardan las coordenadas XYZ en milímetros.

En cada captura se obtienen secuencias de 100 cuadros con las nubes de puntos tridimensionales de cada sensor que son almacenadas inicialmente en los Equipos de Captura y después son transferidos al Equipo de Control para llevar a cabo el resto del procesamiento. En el Equipo de Control las secuencia de nubes de puntos tridimensionales son almacenadas de manera organizada en una base de datos. La Figura 1 muestra la estructura de la base de datos que cuenta con los siguientes niveles: **Nivel 1**.- Se crea una carpeta que lleva en el nombre la palabra “Muestras” seguida de la fecha de la sesión de captura. **Nivel 2**.- Se crea una carpeta para cada secuencia adquirida en la sesión de captura, su nombre inicia con “S” seguido del número de muestra en secuencia de aparición. **Nivel 3**.- Se crea una carpeta por cada Equipo de Captura, *K1* para el sensor 1 y *K2* para el sensor 2. En este mismo nivel se fueron agregando carpetas con los resultados de cada etapa de procesamiento como datos fusionados (FUSEDPC) y datos fusionados normalizados (NFPC). Además de un archivo que contiene la descripción del modelo de movimiento generado.

Se adquieren dos secuencias de datos por cada persona que participa en la sesión de captura. En las secuencias impares (*S1, S3, S5, …*) las personas caminan de derecha a izquierda con respecto a *K1* y en las secuencias pares (*S2, S4, S6, …*) las personas caminan de izquierda a derecha con respecto a *K1*. La base de datos cuenta con 40 secuencias de 20 personas diferentes, 7 mujeres y 13 hombres con edades entre 18 y 50 años, incluyendo un par de gemelos.

2.2. Esquema de Fusión de Datos de Dos Sensores RGBD

Las secuencias de nubes de puntos son procesadas en parejas, nos referimos a cada pareja de nubes de puntos como S_n-$K1_f$ y S_n-$K2_f$, donde $n=0…40$ se refiere al número de secuencia en la base de datos y $f=0…100$ es el número de *frame* o cuadro en la secuencia. Cada pareja de nubes de puntos tridimensionales

será fusionada para obtener una sola nube de puntos denominada $S_nFUSEDPC_f$.

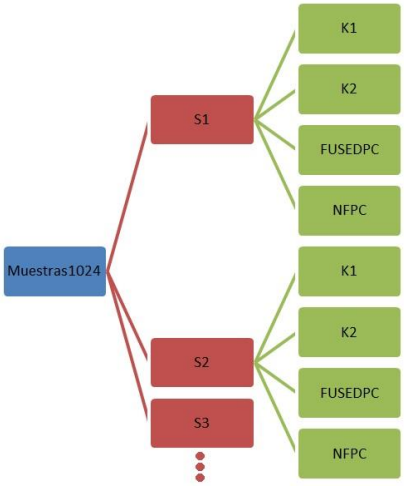
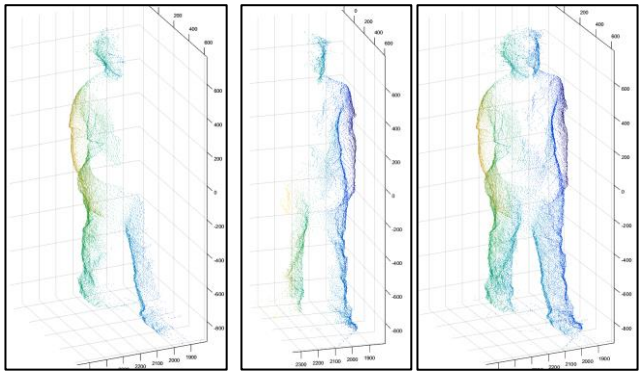


Figura 1.- Estructura de la base de datos de personas caminando.

El proceso de fusión consiste en llevar a cabo una trasformación geométrica sobre la pareja de nubes de puntos S_n-$K1_f$ y S_n-$K2_f$ para alinear sus ejes coordenados con respecto a un mismo punto de referencia y una vez alineados sus ejes proceder a fusionarlas (unirlas) en una sola nube de puntos.

$$S_nFUSEDPC_f = (R1 \bullet S_n\text{-}K1_f + T1)\ U\ (R2 \bullet S_n\text{-}K2_f + T2) \quad (1)$$

En la Ec. 1 podemos ver que la transformación geométrica aplicada a la pareja de nubes de puntos consiste de una rotación *R* seguida de una traslación *T*. Tanto *R1* y *T1*, como *R2* y *T2* son definidas con las distancias y ángulos de rotación necesarios para que las coordenadas del centro de los ejes coordenados de la pareja de nubes de puntos sean transformados y tengan el mismo eje coordenado ubicado en el centro de los dos sensores. La Figura 2 muestra el resultado del proceso de fusión aplicado a una pareja de nubes de puntos de la secuencia S_1 para el cuadro número 20.



(a) (b) (c)

Figura 2.- a) S_1K1_{20}, b) S_1K2_{20} y c) $S_1FUSEDPC_{20}$.

3. MODELO DE LA FORMA DE ANDAR

El objetivo de esta investigación es construir un modelo tridimensional de la forma de andar de una persona con los datos fusionados de dos sensores RGBD, el modelo obtenido podrá será utilizado para el reconocimiento de personas. La hipótesis planteada es que al usar datos de dos sensores capturados desde dos puntos de vista opuestos permitirá reducir las oclusiones y obtener un sistema mejor informado que proporcionará mejores resultados. Para construir un sistema de reconocimiento de personas basado en su forma de andar primero es necesario preparar un modelo representativo que permita una correcta comparación de los datos de diferentes personas. En el caso de los modelos de la forma de andar se requiere de procesos de ajuste y normalización de los datos en las secuencias de nubes puntos fusionados antes de generar el modelo.

3.1. Extracción de un Ciclo Completo de la Forma de Andar

El primer ajuste se refiere a la extracción de un ciclo completo de la forma de andar. Las secuencias de 100 nubes de puntos fusionadas obtenidas de cada persona caminando comprenden normalmente 3 o 4 pasos de los cuales se extrajo un ciclo completo de la forma de andar. Este ciclo completo lo podemos definir como el periodo de tiempo o secuencia de movimientos locomotrices que ocurren desde que uno de los pies tiene contacto con el suelo para iniciar el movimiento del centro de gravedad de la persona hasta que el mismo pie entra de nuevo en contacto con el suelo. Para extraer el ciclo de la forma de andar observamos el ancho de la silueta tridimensional con respecto al eje *X*. Este ancho varía conforme la persona camina, observándose el momento donde el valor del ancho es mínimo cuando los pies de la persona se encuentran juntos y donde el valor del ancho es máximo cuando los pies se encuentran más separados. El objetivo es determinar los puntos máximos y mínimos de la secuencia. Para obtener estos puntos se ajusta un polinomio a los datos del ancho de las siluetas y se calculan las raíces de la derivada del polinomio ajustado. Las raíces obtenidas pueden ser reales o imaginarias, el análisis se realiza seleccionando solo las raíces reales y se determina a que número de cuadro de la secuencia pertenecen.

El inicio del ciclo de la forma de andar se determina con el primer mínimo encontrado y el final del ciclo será ubicado cuatro raíces después. De esta forma se obtiene el número de cuadro correspondiente al inicio y final del ciclo completo de la forma de andar. El modelo de movimiento será generado usando sólo las nubes de puntos fusionadas que corresponden a este ciclo. El número de cuadros correspondientes a un ciclo completo varían según la persona entre 37 y 50 cuadros y es dependiente de la velocidad con la que la persona camina y la amplitud con la que mueve sus extremidades. Este método de extracción de ciclos completos de la forma de andar es una adaptación para datos tridimensionales de la versión en 2D reportada en [8]. La Figura 3 muestra el polinomio ajustado (línea azul) a los datos del ancho de las siluetas (círculos

verdes) y la determinación de los máximos y mínimos (asteriscos rojos) que marcan las diferentes posturas durante un ciclo completo de la forma de andar.

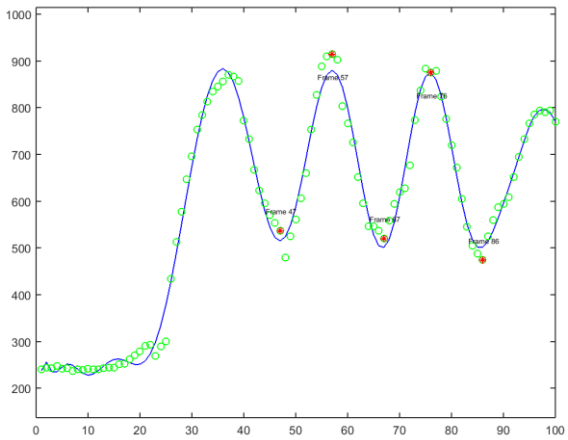


Figura 3.- Gráfica resultante del proceso de extracción de un ciclo completo de la forma de andar.

3.2. Normalización de los Datos Fusionados

Otro ajuste necesario es la normalización de los datos en las nubes de puntos fusionadas. Para llevar a cabo este proceso de normalización se generó un volumen rectangular de un tamaño definido que acomode las nubes de puntos tridimensionales para todas las muestras. El tamaño de este volumen se definió de acuerdo a los máximos posibles en los 3 ejes coordenados de todas las secuencias de nubes de puntos dentro de la base de datos que fueron 1200mm en el eje *X*, 2000mm en el eje *Y* y 1000mm en el eje *Z*. Las nubes de puntos fueron ajustadas para centrar los datos con respecto a los ejes *X* y *Z* de modo que el centro de masa del cuerpo, determinado ubicando el centro de la nube de puntos fusionada, siempre corresponda con el centro del volumen como punto de referencia.

Los datos de las nubes de puntos fusionadas son normalizados en el rango de 0 a 1 y centrados en la coordenada 0.5 con respecto al eje *X*, en el rango de 0 a 1 en el eje *Y* siendo el valor de 0 el nivel del suelo; y en el rango de -0.5 a 0.5 en el eje *Z* centrados en la coordenada 0. La Figura 4 muestra la nube de puntos $S_1FUSEDPC_{20}$ mostrada en la Figura 2 (c) después del proceso de normalización.

Se cuenta con una base de datos de 40 secuencias, en las 20 secuencias impares las personas caminan de derecha a izquierda (de – a + sobre el eje *X*) y en las 20 secuencias pares las personas caminan de izquierda a derecha (de + a – sobre el eje *X*) con respecto al sensor *K1*. Durante el proceso de normalización se invierten las coordenadas en el eje *X* de las secuencias pares de manera que la silueta se mueva en la misma dirección en las 40 secuencias.

Los datos de las nubes de puntos fusionadas y normalizadas son almacenados en la base de datos de acuerdo a la estructura mostrada en la Figura 1.

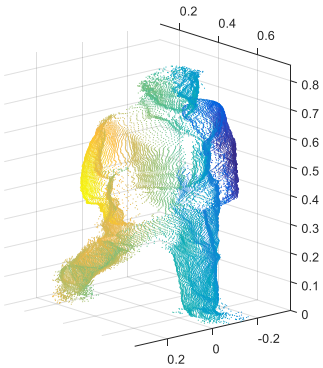


Figura 4.- $S_1FUSEDPC_{20}$ normalizada y centrada.

3.3. Volúmenes de Datos Acumulados (VDA)

Las nubes de puntos de una secuencia normalizada que contienen un ciclo completo de la forma de andar son acumuladas en una sola nube de puntos. La estructura de datos utilizada para acumular las nubes de puntos es un arreglo tridimensional con 100 separaciones por cada eje, es decir, $100x100x100=1x10^6$ casillas, denominado Volumen de Datos Acumulados (VDA). Tomando en cuenta el tamaño de los volúmenes de datos a acumular, esto representa una resolución de 12mm por casilla en el eje *X*, 20mm por casilla en el eje *Y* y 10mm por casilla en el eje *Z*. La cuenta de cada casilla se va incrementando al ir asignando los puntos de cada nube en la secuencia de acuerdo a sus coordenadas.

El VDA resultante contiene información de frecuencias de puntos acumulados de un ciclo de la forma de andar de una persona ya que incluye el movimiento completo del cuerpo donde se observan mayores frecuencias en los puntos que corresponden al tronco y cabeza, y menos frecuencia en las extremidades ya que son las partes del cuerpo que presentan mayor cantidad de movimiento. En la Figura 5 se muestra la visualización de uno de estos VDA como una imagen en donde los puntos más claros representan mayor frecuencia de puntos acumulados.

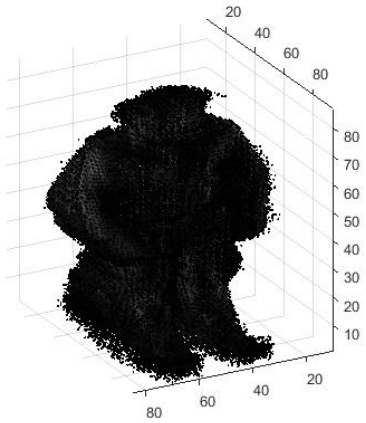


Figura 5.- Volumen de Datos Acumulados (VDA).

Una vez que todos los puntos de los volúmenes del ciclo han sido asignados a su casilla correspondiente en el VDA, este se convierte en un vector con $1x10^6$ elementos y se guarda en un archivo que representa el modelo de la forma de andar de la persona.
La manera en que se construyen los VDA es similar a la descrita en [9], donde utilizan datos de vista frontal de la forma de andar de la persona con un solo sensor RGBD para construir el volumen acumulado, mientras que en este artículo se desarrolló un modelo que utiliza datos de dos sensores colocados en vistas opuestas a los lados de la persona en movimiento.

4. EXPERIMENTO DE RECONOCIMIENTO DE PERSONAS

El experimento está basado en trabajos previos donde se utiliza información de las siluetas en 2D de un ciclo completo de la forma de andar acumuladas en una sola imagen denominada *GEI* (*Gait Energy Image*) [5]. Se realiza un procedimiento similar con los datos tridimensionales para el reconocimiento de personas utilizando la técnica de análisis de componentes principales. Se utilizan los VDA calculados sobre la base de datos de 20 personas para construir un sistema de clasificación de patrones. Se cuenta con dos VDA por cada persona de la base de datos, uno de ellos se utiliza para la etapa de entrenamiento del sistema y el otro para la etapa de prueba.

4.1. Etapa de Entrenamiento

En términos de teoría de información, se pretende extraer la información relevante del modelo de movimiento, codificado en forma de un VDA, de la manera más eficiente posible y comparar esta información entre las distintas muestras sin la necesidad de aplicar criterios de segmentación de extremidades u otras técnicas que se utilizan tradicionalmente en el análisis de movimiento humano. Para esto, el objetivo fue encontrar los componentes principales de la distribución de los modelos tridimensionales, es decir, los eigenvectores de una matriz de covarianza donde los modelos tridimensionales son representados como puntos en un espacio multidimensional denominado eigenespacio.
Los eigenvectores son ordenados tomando en cuenta la variación de un modelo muestra a otro. Este esquema de reconocimiento de personas basado en el análisis de componentes principales es similar al presentado en [10].
La etapa de entrenamiento consistió entonces de calcular los eigenvalores y eigenvectores de una matriz de covarianza L calculada de cómo

$$L= A \cdot A' \qquad (2)$$

Donde A es la matriz construida colocando los VDA en forma vectores columna de $n=1x10^6$ elementos y restándole a cada uno el vector promedio de las $m=20$ muestras seleccionadas para la etapa de entrenamiento.

$$A(i)=VDA(i)-VDA_promedio \qquad (3)$$

Donde $i=1$ *a* 20 es el número de columna de la matriz A. La matriz de covarianza L tiene un tamaño de n^2 por lo que el cálculo de sus eigenvalores y eigenvectores se vuelve una tarea computacionalmente intratable. Un método más factible sería considerar el cálculo de la matriz L como

$$L= A'A \qquad (4)$$

Esto es posible ya que el número de puntos en el eigenespacio (20 muestras) es mucho menor al tamaño del espacio ($1x10^6$), entonces la matriz L tendrá un tamaño de 20x20 y de estas se pueden calcular los 20 eigenvectores más significativos. Estos eigenvectores determinan las combinaciones lineales de las 20 muestras del conjunto de entrenamiento para calcular los eigenvectores $\boldsymbol{u}$ por medio de

$$\boldsymbol{u}=A*V \qquad (5)$$

Donde V es una matriz que tiene como columnas a los 20 eigenvectores de la matriz L.

4.2. Etapa de Prueba

Obtenidos los vectores característicos o eigenvectores más significativos de las muestras de entrenamiento se procede a la etapa de prueba utilizando las muestras reservadas para esta etapa del sistema de reconocimiento. Se cuenta con 20 eigenvectores que forman un conjunto base que describe los modelos de la forma de andar de los 20 sujetos de la base de datos.
Cada uno de los 20 VDA del conjunto de muestras de prueba es transformado en un eigenvector, es decir, es proyectado al eigenespacio por medio de

$$\boldsymbol{w}(i)=\boldsymbol{u}' * (VDA(i)-VDA_promedio) \qquad (6)$$

Los pesos $\boldsymbol{w}(i)$ forman el vector $\boldsymbol{W}$ que describe la contribución de cada eigenvector de **u** en la representación del *VDA(i)*. El vector de pesos $\boldsymbol{W}$ se usa después para realizar la clasificación encontrando a que sujeto pertenece cada uno de los *VDA(i)* del conjunto de muestras de prueba. El método para determinar la clase a la que pertenecen los VDA de prueba es encontrar la clase c que minimice la distancia euclidiana

$$D=||(\boldsymbol{W}-\boldsymbol{W}_{c})||^2 \qquad (7)$$

Donde $\boldsymbol{W}_c$ es el vector de pesos que representa la *c-ésima* clase de las 20 posibles. Así, un VDA representado por su vector de pesos $\boldsymbol{W}$ es clasificado como perteneciente a la clase c cuando la distancia mínima D es encontrada con respecto a $\boldsymbol{W}_c$. Se calcularon las distancias euclidianas entre las proyecciones de cada VDA de las muestras de prueba, contra las proyecciones de los VDA de las muestras de entrenamiento. Se espera que los VDA correspondientes a la

misma persona en las muestras de entrenamiento y de prueba presenten la distancia euclidiana más pequeña. La tabla completa de distancias inter-clases se puede consultar en [11] donde se observan las distancias euclidianas obtenidas para la base de datos existente, se puede observar que la distancia mínima obtenida de cada muestra de prueba corresponde a la clase de la muestra de entrenamiento y que la distancia con el resto de las clases es mucho mayor. En la etapa de prueba las distancias euclidianas mínimas apuntan a la clase correspondiente correcta de manera que el 100% de las muestras de prueba fueron clasificadas correctamente.

4.3. Experimento de Reconocimiento con Datos de un Solo Sensor RGBD

Se realizó otro experimento siguiendo la misma metodología de las secciones 4.1 y 4.2 pero utilizando los datos de profundidad obtenidos de un solo sensor RGBD para comparar los resultados contra los obtenidos utilizando el modelo tridimensional con dos sensores RGBD. El experimento sigue los mismos procedimientos y cálculos pero utiliza las nubes de puntos obtenidas únicamente desde un solo punto de vista.

Se calcularon, normalizaron y acumularon las nubes de puntos utilizando los mismos procedimientos matemáticos pero con los datos obtenidos de un solo sensor. La Figura 6 muestra el VDA resultante de la acumulación de nubes de puntos sobre una de las secuencias de la base de datos para el sensor *K1*.

Siguiendo la metodología se calcularon los eigenvectores de los VDA de entrenamiento y sus proyecciones al eigenespacio. Se compararon las distancias euclidianas con las proyecciones de los VDA de las muestras de prueba. La tabla completa de distancias inter-clases para el experimento de reconocimiento personas con datos de un solo sensor RGBD se puede consultar en [11] donde se observan las distancias euclidianas obtenidas al comparar las muestras de prueba contra las muestras de entrenamiento.

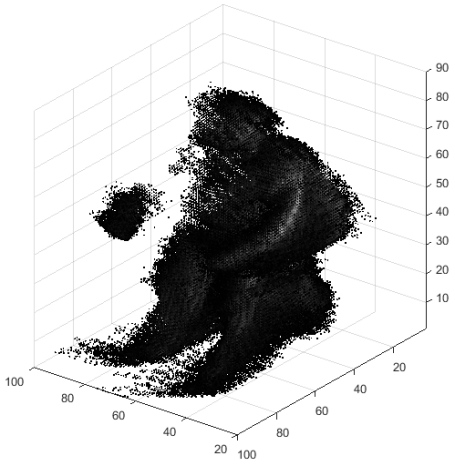


Figura 6.- VDA construido con los datos de un solo sensor RGBD.

El resultado de la etapa de prueba indica que 16 de las 20 muestras de prueba fueron clasificados correctamente, de modo que las muestras 1, 4, 10 y 17 fueron clasificadas de forma incorrecta. El sistema de reconocimiento de personas basado en la forma de andar obtuvo un 80% de certeza en la clasificación con datos de un solo sensor contra un 100% utilizando los dos sensores.

5. CONCLUSIONES

Se desarrolló un modelo que captura el movimiento de una persona, específicamente se modela la forma de andar. Para la construcción del modelo se llevó a cabo la extracción de un ciclo completo de la forma de andar de cada una de las secuencias de nubes de puntos fusionados de la base de datos. Las nubes de puntos correspondientes a un ciclo de la forma de andar fueron normalizadas y centradas para poder hacer una comparación entre diferentes personas. El modelo se construye acumulando las nubes de puntos de un ciclo de la forma de andar en un solo Volumen de Datos Acumulados (VDA) que captura la forma de andar de la persona.

Se tiene una base de datos de 20 personas con dos secuencias por persona, se calcularon los VDA de las 40 secuencias y se construyó un sistema de reconocimiento basado en el modelo de la forma de andar desarrollado. Se utilizó un VDA de cada persona para formar el conjunto de 20 muestras para la etapa de entrenamiento y con las 20 muestras restantes se formó el conjunto de muestras para la etapa de prueba. Se utilizó la técnica de análisis de componentes principales para construir un eigenespacio con el conjunto de muestras de entrenamiento. Después se proyectó cada muestra del conjunto de prueba a este eigenespacio y se calculó la distancia con respecto a cada clase o muestra de entrenamiento, entonces se clasificó la muestra de prueba de acurdo a la menor distancia euclidiana. El sistema de reconocimiento clasificó correctamente el total de las muestras.

Se construyó un sistema de reconocimiento utilizando los datos de un solo sensor RGBD siguiendo el mismo procedimiento descrito para los datos fusionados. Se calcularon los VDA para las secuencias de nubes de puntos de un ciclo de la forma de andar. Se entrenó este sistema con 20 de las muestras y se probó con las 20 restantes, la clasificación no fue correcta para 4 de las muestras. De esta forma concluimos que nuestra hipótesis fue comprobada ya que al tener el sistema mejor informado, con datos adquiridos desde dos puntos de vista opuestos, se tiene mejor desempeño que el sistema construido con los datos de un solo sensor RGBD.

Este artículo es resultado del desarrollo del proyecto "**Procesamiento de Múltiples Flujos de Video RGBD para Análisis de Movimiento Humano**" con clave **5551.15.16-P-P** financiado por el Tecnológico Nacional de México.

REFERENCIAS

[1] Forsyth, D.A., *Computer Vision, a Modern Approach*, 2nd edition, Pearson, 2012.

[2] Philips, P. J., "Human Identification Technical Challenge", *International Conference on Image Processing Proceedings*, pp. 49-50, 2002.

[3] Sarkar, S., Philips, P. J., Liu, Z., Robledo-Vega, I., Grother, P., Bowyer, K.W., "The HumanID Gait Challenge Problem: Data Sets, Performance, and Analysis", *Transactions on Pattern Analysis and Machine Intelligence, Vol. 27, No. 2*, pp. 162-177, 2005.

[4] D.S. Matovski, M.S. Nixon, J.N. Carter, Gait recognition, in: Computer Vision, Springer, 2014, pp. 309–318.

[5] Man, Ju, and Bir Bhanu. "Individual recognition using gait energy image." *IEEE transactions on pattern analysis and machine intelligence* 28.2 (2006): 316-322.

[6] Lulu Chen, Hong Wei, James Ferryman, "A Survey of Human Motion Analysis Using Depth Imagery", *Pattern Recognition Letters*, Vol. 34, pp.1995-2006, Nov. 2013.

[7] Velasco, V.H., Robledo,I., Pacheco, A., "Fusión de Datos de Dos Sensores RGBD para Captura de Movimiento Humano", *38º Congreso Internacional de Ingeniería Electrónica, ELECTRO 2016*, Instituto Tecnológico de Chihuahua, Octubre 2016.

[8] Robledo, I., Gutierrez, D.M., "Procesamiento de Videos de Personas Caminando para la Extracción Automática de un Ciclo Completo de la Forma de Andar", *6º Congreso Internacional de Investigación CIPITECH 2013*, Chihuahua, México, Octubre 2013.

[9] Sivapalan, S., Chen, D., Denman, S., Sridharan, S., Fookes, C., "Gait Energy Volumes and Frontal Gait Recognition using Depth Images", *International Joint Conference on Biometrics*, 2011.

[10] Turk, Matthew, and Alex Pentland. "Eigenfaces for recognition." *Journal of cognitive neuroscience* 3.1 (1991): 71-86.

[11] Velasco, V.H., "Fusión de Datos de Dos Sensores RGBD para Generar un Modelo de 360 Grados del Movimiento de una Persona", *Tesis de Maestría*, Instituto Tecnológico de Chihuahua, 2017.