

Identificación de movimiento humano a través del procesamiento de video¹

Identification of human movement through video processing

Y. T. Leguizamón, W. H. Sánchez y E. L. Sierra

Recibido: julio 27 de 2018 – Aceptado: septiembre 30 de 2019.

Resumen— El estudio del movimiento del cuerpo humano y el reconocimiento del rostro han abierto fronteras en las diferentes áreas del conocimiento. Este artículo tiene como objetivo presentar la generación de un método alternativo para identificar rostros y caracterizar el movimiento del cuerpo humano. La metodología utilizada tiene tres fases, en la primera se genera una base de videos NoSQL haciendo uso del sistema gestor de bases de datos (SGDB) MongoDB; la segunda permite la identificación de rostros humanos mediante dos algoritmos como la Transformada de Hough y el algoritmo de Viola-Jones; y la tercera identificación de movimiento utilizando Flujo Óptico y el método de Histograma de Gradientes Orientados. Como resultados significativos se obtiene que la identificación facial permite seleccionar un rango de relevancia que da las coordenadas del área aproximada de interés o actor, para ser tenido en cuenta al realizar el procesamiento de flujo óptico, evitando así, procesar áreas de falsos positivos, es decir sombras o reflejos producidos por el ambiente. Finalmente se concluye que, al tener una identificación clara del rostro en vista frontal, facilita el seguimiento de las acciones humanas.

Palabras clave— Procesamiento de video, reconocimiento de actividad.

Abstract— Studies of face recognition and movements of the human body have opened frontiers for different knowledge areas. The objective of this paper is to show an alternative methodology to identify faces and make a characterization of the human body movement. This methodology has three phases. The first phase is

a generation of a video database in a NoSQL model using MongoDB. The second phase permits to identify human faces applying an algorithm that combines Hough Transform and the Viola-Jones algorithm. And the third phase makes an identification of human movement using optical flow histogram-oriented gradients. The main results show that the face identification permits to select a relevance range for an approximation of the interest or actor area. This interest area is a base to apply the optical flow, avoiding false positives, like shadows or reflections produced by the environment. Finally, we conclude that the obtainment of a clear frontal view of the face supplies an easy way for human action tracking.

Key words— Video processing, activity recognition.

I. INTRODUCCIÓN

LA multimedia permite hacer uso de herramientas digitales para crear sistemas que logran identificar el comportamiento humano [1]. Para lograr esta identificación de comportamiento se hace énfasis en la comunicación corporal que tienen las personas, por lo que se requiere un estudio del movimiento (kinestesia), comportamiento y de la relación de un ser humano con su entorno, para la posterior clasificación de las acciones que genere. Esta información es vital para la elaboración de cualquier sistema que tenga como meta la identificación y clasificación de acciones humanas.

El estudio del movimiento humano ha abierto diferentes fronteras para el desarrollo en el área del entretenimiento [2]. En el caso de películas como Avatar, El Planeta de los Simios, El Señor de los Anillos, entre otras, y videojuegos como The Last of Us, Beyond: Two Souls, Heavy Rain, etc., se usan actores reales para capturar el movimiento natural de un ser humano, por medio de trajes de captura de movimiento [2]. Los trajes están conformados por sensores en las articulaciones, extremidades y torso, para capturar señales en tiempo real, que crean patrones de movimiento del cuerpo y dan vida a un personaje de Imagen Creada por Computador (CGI) [3]. Este es uno de los métodos más efectivos para la captura de movimiento, pero su adquisición es de carácter limitado, estos trajes son costosos, además se debe tener un mínimo de conocimiento para usarlos, al igual que software y hardware especializado [3].

Existen otros artefactos como el Kinect que detecta la

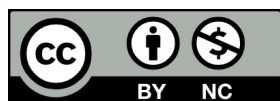
¹Producto de investigación del Grupo de Investigación GIM, Semillero SAMI, de la Universidad Militar Nueva Granada, Facultad de ingeniería.

Y. T. Leguizamón, Universidad Militar Nueva Granada, Bogotá, Colombia, email: u1201582@unimilitar.edu.co.

W. H. Sánchez, Universidad Militar Nueva Granada, Bogotá, Colombia, email: wilman.sanchez@unimilitar.edu.co.

E. L. Sierra, Universidad Militar Nueva Granada, Bogotá, Colombia, email: eduard.sierra@unimilitar.edu.co

Como citar este artículo: Leguizamón, Y. T., Sánchez, W. H., y Sierra, E. L. Identificación de movimiento humano a través del Procesamiento de video, Entre Ciencia e Ingeniería, vol. 13, no. 26, pp. 27-35, julio-diciembre 2019. DOI: <https://doi.org/10.31908/19098367.1152>.



forma de un individuo e identifica su movimiento [4]. Este dispositivo captura el movimiento de la persona sin usar trajes. Este sistema es más asequible para el público, pero tiene limitantes como la necesidad de un gran espacio vacío para interactuar. También, se presenta problema de oclusión, que pueden originar una lectura errónea de movimiento. Como solución, se considera el uso de múltiples vistas con más sensores de Kinect para disminuir el rango de error [4]. Otro problema es el nivel de iluminación, ya que una gran intensidad de luz causa una mala lectura de la figura humana.

Varias personas han optado buscar otro tipo de herramientas y conceptos para lograr una buena captura de movimiento sin hacer uso de los sistemas costosos y complejos anteriormente mencionados. En el artículo denominado “Sistema de seguimiento del cambio de la postura de una persona que realiza una actividad en un lugar cerrado” realiza un proceso de segmentación para identificación de poses en un proceso donde se aplica la detección de bordes de cada imagen, aplicación de niveles de umbralización, operaciones en el imágenes en escala de grises, proceso donde se obtienen puntos de reconocimiento de objetos, de los cuales se hace una clasificación de píxeles, un entrenamiento de estos, y de este resultado se obtienen parámetros para caracterizar la poses de cada una de la imágenes [5].

En el caso de los autores Corazza, Mündermann y Andriacchi se inició el proceso de una identificación de movimiento sin marcadores con la captura de videos desde varias cámaras donde un individuo realiza alguna acción. Después de tener esta información se hace una reconstrucción del escenario 3D enfocándose en la proyección de las siluetas de cada una de las cámaras usadas, estas siluetas son obtenidas de la separación del primer plano y del fondo de cada uno de los frames capturados. Y por último después de este proceso se hace una vinculación de un modelo 3d con el escenario 3d recreado para que este simule el movimiento de las siluetas obtenido [6]. Así mismo estos autores en otro artículo realizan un proceso similar de capturas de videos de un individuo caminado, tomando como parámetro cámaras sincronizadas y calibradas en todos los aspectos de color. Se hace un enfoque en el modelo 3d creado de las secuencias imágenes de las cámaras y también un modelo 3d en una ambientación 2d, y de la información obtenida se hace la reconstrucción de la secuencia de caminado usando un sistema de visión pasivo [7].

Los sistemas anteriormente mencionados tienen sus pros y sus contras, por lo que se abre campo a considerar una nueva alternativa para obtener información. El presente artículo, propone una alternativa para identificar movimiento realizado por personas, para lo cual se identifica primero los actores mediante la identificación del rostro, con lo cual se evita el preprocesamiento, en especial, de eliminación de sombras y objetos, al igual que reconstrucciones 2D y 3D, y así, solamente capturar datos de movimientos válidos provenientes del individuo identificado en el video. Para lo anterior, se crea una base de datos de videos propia, que cumpla con las características necesarias para el desarrollo del proyecto, ya que la mayoría de las bases que se usan para hacer este tipo de estudios proceden de sujetos de prueba de otras latitudes, lo que implica sujetos con morfología y movimientos sui generis. Cabe aclarar que esta alternativa de identificación de

movimiento se aplicará evitando el uso de trajes y sensores especializados, entre otros.

El artículo se compone de tres secciones, en la segunda se explican los materiales y métodos que se usaron para el desarrollo del proyecto, la tercera sección muestra los resultados y el análisis de los datos, y por último se finaliza con conclusiones obtenidas.

II. MATERIALES Y MÉTODOS

El artículo se divide en tres áreas principales, la primera habla sobre la creación de la base de datos de videos según movimiento, en la segunda área se aplican algoritmos de reconocimiento de rostros para identificar personajes en los videos de la base de datos y la tercera área trata la caracterización de movimiento de personas identificadas en los videos (Fig. 1).

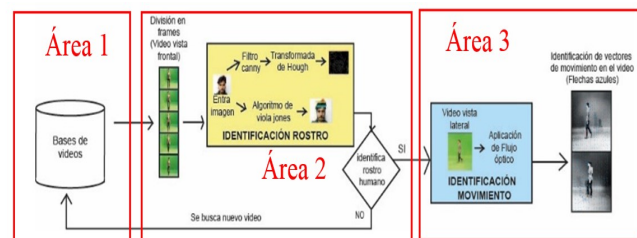


Fig. 1. Esquema de funcionamiento.

A. Área # 1 Bases De Datos De Videos

Como primera instancia se procede a crear una base de datos NoSQL haciendo uso del sistema gestor de bases de datos (SGDB) MongoDB, debido a que este permite un almacenamiento de gran cantidad de videos sin perder velocidad de funcionamiento. La base de datos implementada es poblada por una serie de videos que contienen grabaciones de movimiento humano usando diferentes parámetros de clasificación propias del video como grupo de movimiento, subgrupo de animación, poses de animación, número de secuencia y tipo de plano, asimismo parámetros de las personas que participaron en cada uno de los videos como, género, nombre, población y tipo somático.

Cada video cumple con una serie de características para el proyecto: se hace la toma de videos en un ambiente con croma de fondo, con una iluminación general y el uso de cámaras profesionales. En cada uno de los videos hay inmerso un movimiento específico ejecutado por actores de prueba. Los videos tienen una duración máxima de 7 segundos y se toman desde un plano frontal y lateral del actor. El plano frontal se usa para la identificación inicial de rostros de los individuos en cada uno de los videos, mientras que el plano lateral se aplica para la identificación de la acción que esté realizando el individuo en el video.

Para determinar qué movimientos se van a trabajar en el proceso se crea una tabla de planeamiento de grabación donde se tienen tres secciones principales: la descripción que contiene 3 grandes grupos de movimiento que son movimientos normales del ser humano; enfocados netamente a animación y movimientos especializados en una disciplina, junto con las secuencias y poses que representan los

subgrupos que componen los movimientos principales y el orden, que es un identificador inicial que se da a cada subgrupo para clasificar los videos. Los grupos se crean según el nivel de complejidad de la acción que se tiene que realizar; para trabajar en primera instancia se opta por el grupo de movimientos normales del ser humano, debido a que son acciones destacables y comunes de cambio de poses y con ellas se obtiene el primer paso para entrenar el sistema para los posteriores grupos y acciones que se vayan a trabajar (Tabla I).

TABLA I
SECCIÓN DE PLANEACIÓN DE SECUENCIAS DE ANIMACIÓN

Descripción	Secuencias - Poses de Animación	Orden
M 1 Movimientos normales del ser humano.	Movimientos cotidianos	G1
	Movimientos básicos generales del cuerpo humano.	G2
	Movimiento por estiramientos básicos.	G3
	Movimientos básicos de calentamiento previo a ejercicios.	G4

TABLA II
SECCIÓN DE PLANEACIÓN DE GRABACIÓN DE SECUENCIAS DE ANIMACIÓN G1

Numero de secuencia	Orden de grabación	Angulación cámara	Duración del video	Consecutivo
A00	Saludo de manos	Lateral/Frontal	6 seg	001
A05	Caminado	Lateral/Frontal	6 seg	006
A06	Salto	Lateral/Frontal	6 seg	007
A08	Pararse	Lateral/Frontal	6 seg	009
A09	Arrodillarse	Lateral/Frontal	6 seg	010

TABLA III
SECCIÓN DEL METADATO GENERAL DE SECUENCIA DE VIDEO

	Contenido	Propiedad intelectual	
Título	Nombre del recurso	Autor	Persona que dirige la grabación
Descripción	Descripción verbal detallada del recurso	Publicado por	Semillero de investigación
Duración	Tiempo de duración cada video	Derechos	UMNG
Tipo somático	Endomorfo Ectomorfo Mesomorfo	Actor	Identificador del actor
Género	Género del actor	Colaborador	Personas que participan

Después de esta clasificación se crea la tabla que corresponde al subgrupo “Movimientos cotidianos” perteneciente al grupo “Movimientos normales del ser humano denominado”; en ella se puede ver algunas de las secuencias de movimiento que fueron puestas a grabación

(Tabla II).

Al determinar qué subgrupo se va a trabajar para grabar, se crea un metadato general de secuencia de video, en donde describen todos los parámetros mencionados anteriormente para asociar a cada video un identificador único (Tabla III).

B. Área # 2 Identificación De Rostros Humanos

Después de completar el anterior proceso, se procede a pasar a la segunda parte del sistema, que trata de la identificación de rostros en los videos. Esta identificación se hace con la meta de corroborar que en cada uno de los videos de la base creada se encuentra una persona, por lo que es viable su estudio de movimiento, si no se identifica el rostro de un ser humano en el video, el video es descartado por falta de información.

Para este procedimiento, se hace un estudio sobre qué tipos de algoritmos pueden usarse para identificar tanto formas como partes del rostro humano, de los cuales se escogieron dos algoritmos para experimentar: Transformada de Hough [8] y el algoritmo de Viola-Jones [9]. La elección de la transformada Hough se dio por su enfoque a la identificación de figuras, a lo que se desea aplicar en la detección de formas geométricas inmersas en partes del rostro humano; mientras que la elección del algoritmo de Viola-Jones se dio por su alto nivel de asertividad en detección de rostros utilizando parámetros como la ley de tercios, distancias entre otras cosas.

El primer algoritmo considerado es la transformada de Hough. Para el uso de la transformada de Hough se debe tener un control sobre las divisiones del rostro humano, ya que este algoritmo se especializa en encontrar principalmente figuras geométricas en una imagen de forma general [8]. Para probar este algoritmo se hace una base de archivos con ilustraciones en plano frontal, a las cuales se hace un estudio del rostro humano.

El rostro humano se puede dividir en varias secciones simétricas tanto vertical como horizontalmente. Para esta división se tiene en cuenta la ley de tercios, en la cual se habla de dividir la imagen en proporciones que equilibren la imagen. Al hacer este tipo de segmentación las partes que componen la cara se ven separadas por una malla imaginaria como se puede observar en la Fig. 2.

Además de tener este control de divisiones, el algoritmo de Hough permite el control de radio y sensibilidad, que ayudan a ajustar y detectar nuevos elementos [8].

El segundo algoritmo que se considera es el Algoritmo de Viola-Jones, es el primer algoritmo creado para la detección de objetos en tiempo real creado por Paul Viola y Michael Jones [9]. Este se puede usar para la identificación de diferentes elementos, pero fue creado principalmente para la detección de rostros humanos [9] (Fig. 3).

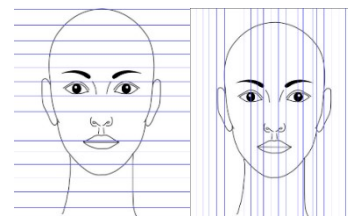


Fig. 2. División del rostro humano por una gravilla en sentido vertical y horizontal.



Fig. 3. Ilustración rango de clasificación con relación en rostro.

Para esta área de identificación se usó una base de archivos de ilustraciones y otra base de rostros humanos propia como fase de prueba, para que una vez que son probados los algoritmos, estos se aplican a los frames extraídos de los videos en vista frontal de la base de datos creada en la primera área (Fig. 4 y Fig. 5).



Fig. 4. Algunas imágenes de la base de archivos de rostros humanos.



Fig. 5. Algunas imágenes de la base de archivos de ilustraciones.

C. Área # 3 Identificación De Movimiento

Al terminar la segunda parte en la cual se identifica la presencia de una persona en los videos por medio de su rostro, se procede a la caracterización e identificación de poses del cuerpo humano.

Para este proceso se aplican dos métodos de identificación de movimiento, los cuales se comparan según las características que se extraen de los videos, y así determinar cuál es más efectivo en detectar cambios de movimiento, estos métodos son: Flujo Óptico, el cual se trabaja directamente en el proyecto y el método de Histograma de Gradientes Orientados que es trabajado en un proyecto en construcción llamado “Identificación de poses del cuerpo humano para obtener información de movimiento.”

El Flujo Óptico es un método que implementa el algoritmo de Lucas kanade, el cual tiene dos algoritmos de ejecución, el primero busca cambios de los niveles de grises en cada uno de los pixeles de cada fotograma del video, los cuales van creando un patrón que es identificado por un segundo algoritmo, que aplica la teoría de vectores de velocidad para graficar la dirección en que los niveles de gris están cambiando, lo cual se asocia a un movimiento identificado [10].

III. RESULTADOS

A. Resultado Área # 1 Bases De Datos De Videos

Al tener las tablas 1,2 y 3 completas se procedió a la toma de grabaciones. Se graban un total de 15 actores, en un ambiente controlado (fondo verde) con diferentes focos de luz. En esta primera sesión de grabación se eligieron los planos

frontal y lateral de cada movimiento de los actores, donde se encuentran las poses que se muestran en la Tabla IV.

TABLA IV
POSES EN FRAMES DE VIDEO

		
Saludo de mano	Caminando	Pararse
		
Salto	Arrodillarse	

B. Resultado Área # 2 Identificación De Rostros Humanos

Como se menciona anteriormente, se hace la aplicación de los algoritmos de identificación de rostro primero en las bases de imágenes de prueba. En el caso de la transformada de Hough, al tener identificadas las divisiones, se procede a aplicar el algoritmo de Hough en las ilustraciones de prueba para verificar su eficacia de identificación. Para este proceso se aplica un filtro, en este caso Canny o Sobel, para hacer lectura de bordes y dejar las características principales de rostro. Estos dos filtros se escogieron tanto por la buena lectura que hacen del contorno en las imágenes, así mismo que facilitan la lectura de información de la transformada de Hough (Fig. 6 y Fig. 7).

Después de la lectura de bordes se aplica el algoritmo de Hough, y para identificar qué partes de la imagen fueron detectadas, se diferencian por colores según su ubicación (Fig. 8).

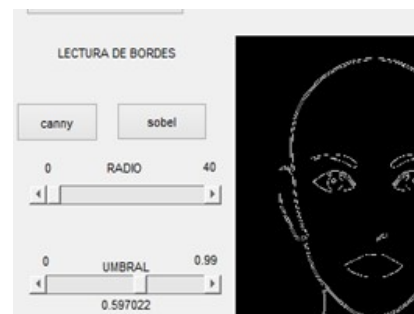


Fig. 6. Ilustración con filtro Canny.

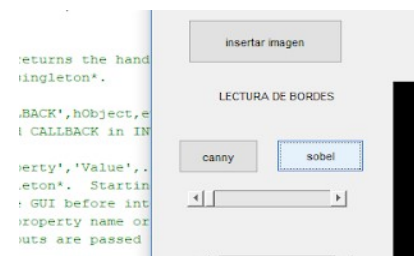


Fig. 7. Ilustración con filtro Sobel.

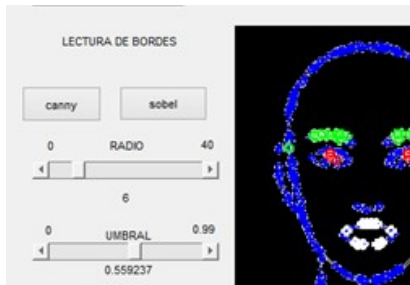


Fig. 8. Identificación de rostros con transformada de Hough en ilustración.

En el caso de Viola Jones, la implantación del algoritmo se aplicó en tres casos usando las bases de archivos: imágenes con filtro Canny, imágenes sin filtro, e imágenes en negativos. En cuanto a las imágenes en negativo solo se trabajó para la base de rostros humanos, ya que las ilustraciones son imágenes en blanco y negro.

El primer caso es imágenes con filtro Canny, en el cual se toman las dos bases de archivos y se les aplica este filtro (Fig. 9 y Fig. 10).



Fig. 9. Algunas imágenes de la base de archivos de rostros humanos con filtro Canny.



Fig. 10. Algunas imágenes de la base de archivos de ilustraciones con filtro Canny.

La aplicación del algoritmo sobre una imagen con filtro Canny da como resultado la imagen que se observa en la Fig. 11.

El segundo caso es imágenes sin filtro, caso en el cual se procede a la aplicación del algoritmo en las imágenes sin modificar. Se hace el procedimiento primero en la base de archivos de rostro (Fig. 12).



Fig. 11. Aplicación de Viola - Jones en imagen de rostro con filtro Canny.

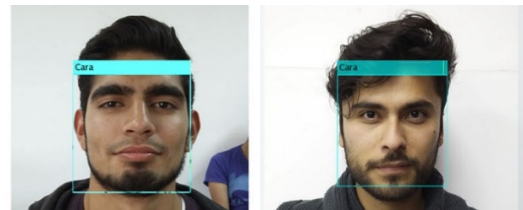


Fig. 12. Aplicación de Viola - Jones en imágenes de rostro humano.

Después se hace el procedimiento en la base de archivos de ilustraciones (Fig. 13).

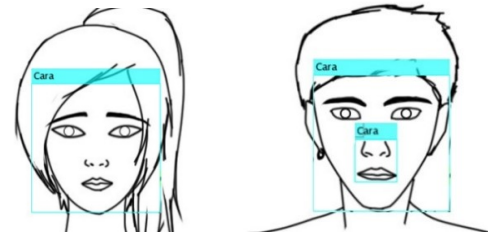


Fig. 13. Aplicación de Viola - Jones en ilustraciones.

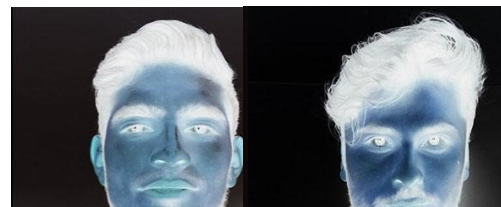


Fig. 14. Aplicación de Viola - Jones en imágenes de rostro en negativo.

El último y tercer caso es imágenes en negativo, caso en que se hace la aplicación del algoritmo a imágenes en negativo de los rostros humanos, que dan el siguiente resultado (Fig. 14).

C. Resultado Área # 3 Identificación De Movimiento

Se aplica el método de Flujo Óptico para esta área. Se hace una primera prueba con una animación de un personaje 2D en un entorno completamente neutro (fondo blanco) (Fig. 15), para tener una mejor captura de información útil que permita minimizar errores de medición.

Después de esta prueba, se usó uno de los videos generados para el proyecto, al cual se le evaluó cambios de niveles de intensidad con el método de Flujo Óptico junto a la identificación de rostros. Como resultado de los dos procesos juntos se obtiene una visualización de flechas o vectores azules mostrando los cambios de movimiento en el video y dos recuadros de diferente tamaño: el recuadro rojo representa la identificación del rostro del individuo y recuadro verde delimita las coordenadas donde se presume que se encuentra la persona en el frame (Fig. 16).

En este sistema se aplica una matriz de $30 * 30$ para definir la dimensión del campo de la velocidad por frame. Esta matriz está compuesta por las coordenadas de velocidad u y v respectivas a las posiciones x y y , coordenadas que van cambiando según la presencia de cambio de nivel de

intensidad en los frames estudiados. Mientras que los componentes x y y tienen una dimensión de $1*30$ y estas especifican las posiciones en la imagen para el trazado de las vectores de velocidad que se encuentren en el frame (Tabla V). El sistema muestra vectores de velocidad con componentes (u, v) en los puntos (x, y) .



Fig. 15. Aplicación Flujo Óptico a animación fondo blanco.

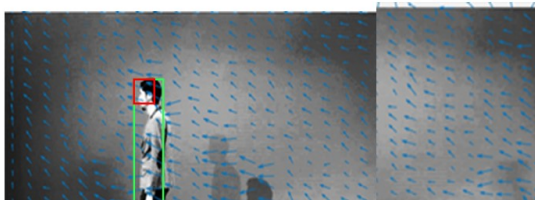


Fig. 16. Aplicación Flujo Óptico a secuencias de imágenes de video (CAMINADO).

El sistema identifica las diferentes partes de los frames que tuvieron un cambio considerable en su tonalidad con respecto a su posición en (x, y) (Tabla VI y Tabla VII). Los valores resaltados en las tablas representan los cambios detectados de movimiento de frame a frame, lo que indica la detección de la acción del individuo en el video.

TABLA V
IDENTIFICACIÓN DE CAMBIOS DE GRISES EN POSICIÓN X Y EN UN FRAME

Posición X	Posición Y
12	393
36	379
60	365
...	...
660	15
684	1
684	1

El sistema procesa cuales son los componentes u y v que se grafican en la secuencia completa de video. En los anteriores pequeños fragmentos se puede observar el cambio de velocidad que se detecta en los pixeles de un frame a otro, el cual se usa como parámetro para dibujar y darle una dirección a las flechas que simbolizan los vectores de movimiento en el video.

TABLA VI

FRAGMENTO DE COLUMNA DE 23 A 28- COMPONENTE U CON DATOS DE VELOCIDAD IDENTIFICADOS EN EL VIDEO DE ENTRENAMIENTO

-0.5392	-0.2154	0.1507	-0.5392	-0.5392	-0.2946	-0.1009
-0.3521	-0.2044	0.0739	-0.3973	-0.5392	-0.1555	-0.1732
-0.3476	-0.3084	-0.5392	-0.2933	-0.5392	-0.2032	-0.1385
-0.3492	-0.3096	-0.5392	-0.5392	-0.5392	-0.2167	-0.0494
-0.3055	-0.3210	-0.5392	-0.5392	-0.5392	-0.1907	-0.1013

TABLA VII

FRAGMENTO DE COLUMNA DE 23 A 28- COMPONENTE V CON DATOS DE VELOCIDAD IDENTIFICADOS EN EL VIDEO DE ENTRENAMIENTO

-0.3690	-0.3690	0.2410	-0.1673	-0.2104	-0.1847	-0.1592
-0.3690	-0.3690	0.2172	-0.1276	-0.2372	-0.0809	-0.1230
-0.3690	-0.3690	0.2755	-0.1104	-0.0520	-0.0565	-0.0507
0.0601	-0.0907	0.3690	-0.2582	-0.2167	-0.0038	-0.1416
0.0187	-0.3690	0.2445	-0.1479	-0.0318	-0.0988	-0.0897

Se hace una nueva lectura de otros vídeos de prueba haciendo nuevas acciones de las cuales se extraen información relevante (Fig. 17 y Fig. 18) (Tabla VIII, Tabla IX, Tabla X y Tabla XI).

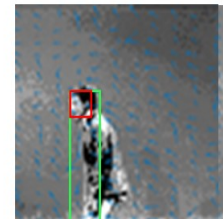


Fig. 17. Aplicación Flujo Óptico a secuencias de imágenes de video (AGACHARSE Y LEVANTARSE).

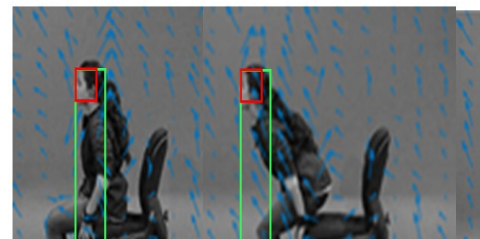


Fig. 18. Aplicación Flujo Óptico a secuencias de imágenes de video (SENTARSE).

TABLA VIII

FRAGMENTO DE COLUMNA DE 1 A 11- COMPONENTE U CON DATOS DE VELOCIDAD IDENTIFICADOS EN EL VIDEO DE ENTRENAMIENTO

0.0151	-0.1511	-0.1335	-0.1744	-0.1131	-0.1439	-0.1317
0.0035	-0.0866	-0.0822	-0.0805	-0.0670	-0.2000	-0.1294
-0.0055	-0.0960	-0.0603	-0.0045	0.0088	-0.0650	-0.0406
-0.0061	-0.0964	-0.0686	-0.0229	0.0001	-0.0742	-0.0348
-0.0039	-0.1011	-0.1185	-0.0504	-0.0763	-0.0654	-0.0323

TABLA IX
FRAGMENTO DE COLUMNA DE I A 11- COMPONENTE V CON DATOS DE VELOCIDAD IDENTIFICADOS EN EL VIDEO DE ENTRENAMIENTO

-0.1368	0.0264	-0.0815	-0.1055	-0.0958	-0.0930	-0.0860
-0.1315	0.0253	-0.0793	-0.1021	-0.0947	-0.0930	-0.0831
-0.0022	-0.0003	-0.0145	-0.0218	-0.0447	-0.0192	-0.0890
-0.0362	0.0012	-0.0697	-0.0056	-0.0240	-0.0223	-0.0903
-0.0314	-0.0138	-0.0434	-0.0451	-0.0383	-0.0027	-0.0518

TABLA X
FRAGMENTO DE COLUMNA DE I A 11- COMPONENTE U CON DATOS DE VELOCIDAD IDENTIFICADOS EN EL VIDEO DE ENTRENAMIENTO

0.2219	-0.8815	-0.1708	-0.5560	0.2374	-0.0369	-0.4002
0.0909	-0.3676	-0.1593	-0.7080	0.0822	-0.1005	-0.4051
-0.1985	-0.4023	-0.2891	-0.3820	-0.3451	-0.6016	-0.0535

TABLA XI
FRAGMENTO DE COLUMNA DE I A 11- COMPONENTE V CON DATOS DE VELOCIDAD IDENTIFICADOS EN EL VIDEO DE ENTRENAMIENTO FIG

-0.0520	0.0159	-0.1193	-0.0524	-0.0675	-0.0533	0.0455
-0.0157	0.0038	-0.1020	-0.0363	-0.0308	-0.1112	0.0090
0.0008	-0.0607	-0.0503	-0.0429	-0.0150	-0.0631	-0.0399

IV. ANÁLISIS DE RESULTADOS

A. Área # 1 Bases De Datos De Videos

Los videos obtenidos de la sesión de grabación tuvieron variaciones de luces, creando diferentes marcas de sombras que sirven para probar el nivel de eficacia que tiene la identificación de rostros para no tomar datos provenientes del movimiento de las sombras del individuo (Tabla IV).

El uso de sistema gestor de bases de datos utilizado, MongoDB, ayuda a almacenar de manera óptima los videos recopilados, ya que entre estos no se crea ningún tipo de relación en su almacenaje, por lo que la cantidad de videos que se pueden almacenar es considerablemente grande, sin perder la velocidad de funcionamiento del proceso.

B. Área # 2 Identificación De Rostros Humanos

En la identificación del rostro humano con transformada de Hough, al hacer pruebas de los filtros de Sobel y de Canny, con diferentes imágenes se opta por usar solo éste último, debido a que el filtro Sobel tiende a eliminar o cortar muchas de las curvas necesarias para identificar correctamente un rostro por Hough como se observa en la Fig. 7.

Al observar el área de resultados, en la Fig. 8, se obtuvo una identificación completa de un rostro, al comprobar que las partes de ojos, nariz, boca y orejas se encontraban inmersas en las dimensiones correspondientes del rostro, por lo que el nivel del umbral usado fue óptimo para tal identificación.

Debido a los parámetros preestablecidos de división de rostro, se logra hacer una identificación total de partes del

rostro, en primera instancia con la transformada de Hough aproximando el rostro y sus partes a círculos.

En la identificación del rostro humano con Viola Jones, al momento de aplicar el algoritmo en una imagen binomial, (imagen con filtro Canny) los clasificadores no pueden tomar cuadros específicos de la imagen, ya que esta se restringe a dos intensidades de pixeles: negro (0) y blanco (255), por lo cual la identificación por secciones no fue posible, como se observa en la Fig. 11.

En el caso de usar la detección en imágenes a color (Fig. 12) el algoritmo logro una buena identificación en la mayoría de las imágenes que fueron puestas a prueba. En el caso de las ilustraciones, el algoritmo identificó más de una vez un rostro en algunas imágenes, mientras que en otras hizo una identificación correcta (Fig. 13). La doble identificación puede deberse a la forma que el rostro está ilustrado, siendo más específico a la separación de ojos, nariz y boca, debido a que cuando se detectaba otro rostro en la ilustración esta identificación siempre se situaba en el área de boca y nariz.

En el caso de las imágenes en negativo (Fig. 14), se puede observar que el algoritmo no pudo lograr una identificación concreta.

Después de terminar el proceso de los tres tipos de imagen evaluados, se observó que el uso de filtros para la detección de bordes así mismo como cambio de valores a negativo no son siempre efectivos para todos los tipos de algoritmos de identificación de objetos, debido a que cada uno toma diferentes valores de intensidades para hacer su respectiva segmentación y clasificación de información.

Para definir que algoritmo da el mejor resultado en identificación, se hace una comparación del porcentaje de acierto de los dos métodos trabajados con las dos bases de archivos trabajadas. En el caso de la transformada de Hough, de un total de 30 imágenes, 12 fueron identificadas de manera correcta, lo que representa un 40% de acierto. Mientras que en el caso del algoritmo de Viola Jones hubo una identificación correcta de 21 imágenes sobre las 30 evaluadas, lo que corresponde al 70 % de acierto.

Después de hacer esta comparación de los métodos de identificación, se opta por el de Viola Jones, debido a su mejor nivel de acierto, así mismo que es más adaptable a situaciones donde las imágenes no se pueden controlar por regiones sumamente estrictas.

C. Área # 3 Identificación De Movimiento

En las Figs. 16 a 18 se muestra la experimentación con el Flujo Óptico, en las que se grafican diferentes vectores que indican la dirección del Flujo en que los pixeles van cambiando, creando una figura básica del movimiento del personaje. Pero al capturar una gran cantidad de información no se pueden obtener específicamente puntos clave como coordenadas para crear una curva clara de animación, por lo que su procesamiento tiene muchos puntos que se podrían interpretar como claves, lo cual ralentizará todo el proceso.

Por otro lado, se procede a hacer la identificación facial mientras se hace estudio de movimiento, ya que el Flujo Óptico hace captura del movimiento de las sombras producidas por los actores u objetos de la escena, creando

falsos positivos de movimiento humano, lo cual afecta el nivel de asertividad de lectura.

La identificación facial permite seleccionar un rango de relevancia que da las coordenadas del área aproximada de interés o actor, para ser tenido en cuenta al realizar el procesamiento de flujo óptico, evitando así, procesar áreas de falsos positivos, es decir sombras o reflejos producidos por el ambiente.

Mientras se hace la comparación del método de histograma de gradientes orientados, en este se realizó una eliminación de sombras y se hizo una lectura de bordes dejando solo la silueta del personaje. En este proceso se usan una serie de descriptores que se van posicionando según van cambiando los frames, creando así el patrón de movimiento, un patrón claro y específico del contorno del individuo a través del video.

Al comparar el proceso que se llevó a cabo en los dos métodos estudiados, se observa que el método de identificación de rostros en conjunto con flujo óptico dio resultados más rápidos y en un carácter de tiempo real. Este logra crear áreas de intereses muy específicas con respecto a la detección del rostro del individuo minimizando lectura de información de más sin la necesidad de un preprocesamiento para eliminar imperfecciones en el ambiente, mientras que la aplicación que se le dio al histograma de gradientes orientados se necesitó un preprocesamiento que hace el proceso de identificación más lento, ya que para este proceso se necesita una manipulación más técnica.

V. CONCLUSIONES

Al hacer un análisis del desarrollo y resultados que brinda el proyecto, se pueden constatar la presencia de diferentes parámetros y medidas que hay que tener en cuenta para una buena captura de información e identificación de la información que se tiene como objetivo caracterizar. Como primera instancia se debe tener control en la limitación de captura errónea de datos según las características del entorno usando herramientas de bajo costo de funcionamiento y de carácter efectivo para cualquiera de los casos que se pueda presentar.

Con relación a los videos, su duración debe ser tan corta como la acción que se realiza, evitando de esta forma sobrecargar la máquina en el análisis, preprocesamiento y procesamiento del video. Dado el contexto de los videos, no se requiere de una alta resolución de los mismos, por lo que se recomienda usar formatos livianos. Otro dato a tener en cuenta es que para la identificación de personajes y de acciones en un video de ambiente controlado, tan solo se requieren de dos planos: uno frontal y uno lateral, el frontal permite identificar el rostro, y el lateral identifica la acción. Al tener una identificación clara del rostro en el plano frontal, se tiene en cuenta este parámetro para la lectura de movimiento en la vista lateral y así se captura información relevante y acorde a las acciones que pertenecen solo a la persona en el video, descartando sombras y otros elementos que puedan generar algún patrón de movimiento.

En el caso de la identificación de rostros, el mejor método para hacer la caracterización fue el algoritmo de Viola Jones, ya que este no tiene limitantes tan estrictas para poder tomar

datos relevantes que ayuden a la lectura correcta de un rostro humano. Aunque Viola Jones sea un algoritmo de identificación de rostros por defecto, otro tipo de herramientas pueden ser programados para cumplir con una identificación similar, en este caso con el uso de la transformada Hough se obtuvieron resultados positivos al hacer un análisis de figuras geométricas inmersas en el rostro.

Como último parámetro se concluye que, para evitar la lectura de falsos positivos de movimiento en el video, se puede trabajar directamente en el video grabado y hacer un proceso de identificación de rostros, con el fin de evitar un pre procesamiento en el que se haga una eliminación de sombras y nivelación de color para haya una lectura más limpia, con lo cual se identifican las áreas de interés y se realiza una identificación más rápida de las acciones o poses inmersas en el video.

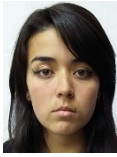
VI. AGRADECIMIENTOS

Este trabajo es financiado con recursos de la Vicerrectoría de Investigaciones de la Universidad Militar Nueva Granada, en el proyecto INV-ING-2617, "Generación de rostro de un personaje a partir de texto". Los autores agradecen el apoyo brindado por la UMNG.

REFERENCIAS

- [1] Manovich, L. El lenguaje de los nuevos medios de comunicación. Cambridge, EEUU: the MIT Press. Available in: < https://books.google.com.co/books?id=jXK7R0hh_aYC&printsec=frontcover&hl=es&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false > Access Date: 18/05/2017, 2001.
- [2] Portafolio. Los fabricantes de videojuegos le apuestan al movimiento y 3D. Available in: <http://www.portafolio.co/economia/finanzas/fabOricante-s-videojuegos-le-apuestan-movimiento-3d-478406> Access Date: 18/05/2017, 2010.
- [3] Fujimori, Y., Ohmura, Y., y Harada, T. Wearable motion capture suit with full-body tactile sensors. *IEEE International Conference on Robotics and Automation..* Available in: <https://ieeexplore-ieee-org.ezproxy.umng.edu.co/document/5152758/> Access Date: 18/05/2017, 2009. DOI:10.1109/robot.2009.5152758.
- [4] Ros, A., y Mendonça, I. Captura de movimiento de personas con múltiples sensores Kinect. Universidad Simon Bolívar. Available in: <http://159.90.80.55/tesis/000155335.pdf> Access Date: 18/05/2017, 2012.
- [5] Buraglia, M. Sistema de seguimiento del cambio de la postura de una persona que realiza una actividad en un lugar cerrado. Available in: <https://repository.javeriana.edu.co/bitstream/handle/10554/7019/tesis466.pdf?sequence=1>. Access Date: 26/06/2018, 2010.
- [6] Corazza, S., Mündermann, L., Andriacchi, M. Motion Capture Methods for the Estimation of Human Body Kinematics. Available in: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.97.1772&rep=rep1&type=pdf> Access Date: 26/06/2018, 2006. doi=10.1.1.97.1772.
- [7] Corazza, S., Mündermann, L. Andriacchi, T. P. The evolution of methods for the capture of human movement leading to markerless motion capture for biomechanical applications. Available in: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1513229/> Access Date: 26/06/2018, 2006. Doi: 10.1186/1743-0003-3-6
- [8] Hough, V., y Paul, C. A. Method and means for recognizing complex patterns. Available in: <http://www.freepatentsonline.com/3069654.html> > Access Date: 03/06/2017, 1962.
- [9] Viola, P., and Jones, M. J. Robust Real-Time Face Detection. *International Journal of Computer Vision.* Available in: < <https://link.springer.com/article/10.1023%2FB%3AVISI.0000013087.49260.fb> Access Date: 03/06/2017, 2004. DOI:10.1023/B:VISI.0000013087.49260.fb

- [10] Lucas, B. D., and Kanade, T. An iterative image-registration technique with an application to stereo vision. *The 7th international joint conference on Artificial intelligence*. doi=10.1.1.421.4619. Available in: <http://citeseer.ist.psu.edu/viewdoc/summary?doi=10.1.1.421.4619>>Access Date: 04/06/2017, 1981.



Yessica Tatiana Leguizamón. Estudiante de décimo semestre de Ingeniería en Multimedia de la Universidad Militar Nueva Granada, Sede Bogotá. Graduada del Colegio Minuto de Dios. Con intereses en el área de procesamiento de imágenes, animación, modelado 3D, dibujo, computación gráfica y render. ORCID: <http://orcid.org/0000-0002-9562-9762>.



8007-1877.

Wilman Helioth Sánchez. Ingeniero de Sistemas de la Universidad Nacional de Colombia. Magister en Educación y Especialista en Docencia Universitaria de la Universidad Militar Nueva Granada. Igualmente ha realizado educación no formal como Auditor Interno de Calidad y posee formación como coordinador de autoevaluación y par académico en procesos de calidad. ORCID: <http://orcid.org/0000-0001-8007-1877>.



<http://orcid.org/0000-0002-0319-6227>.

Eduard Leonardo Sierra-Ballen. Ingeniero de Sistemas de la Universidad Nacional de Colombia, Magister en Educación de la Universidad Militar Nueva Granada. Se encuentra vinculado como profesor Y director en el Programa de Ingeniería en Multimedia, de la misma Universidad, Sede Campus Nueva Granada, Cajicá. Cofundador del Grupo de Investigación en Multimedia GIM. ORCID: <http://orcid.org/0000-0002-0319-6227>.