



CENATAV

Centro de Aplicaciones de
Tecnologías de Avanzada
MINISTERIO DE LA INDUSTRIA BÁSICA

RNPS No. 2142
ISSN 2072-6287
Versión Digital

REPORTE TÉCNICO
**Reconocimiento
de Patrones**

SERIE AZUL

**Seguimiento y conteo de personas
en ambientes exteriores con una
cámara fija**

Lic. Walter Izquierdo Guerra,
Dr. C. Edel García Reyes

RT_018

Diciembre 2009





CENATAV

Centro de Aplicaciones de
Tecnologías de Avanzada
MINISTERIO DE LA INDUSTRIA BÁSICA

RNPS No. 2142
ISSN 2072-6287
Versión Digital

SERIE AZUL

REPORTE TÉCNICO
**Reconocimiento
de Patrones**

**Seguimiento y conteo de personas
en ambientes exteriores con una
cámara fija**

Lic. Walter Izquierdo Guerra,
Dr. C. Edel García Reyes

RT_018

Diciembre 2009



Índice

1.	Seguimiento.....	2
1.1	Introducción.....	2
1.2	Seguimiento basado en modelos.....	3
1.2.1	Modelos tradicionales utilizados para el cuerpo humano.....	4
1.2.1.1	Figura esquelética.....	4
1.2.1.2	Contornos en dos dimensiones (2D).....	5
1.2.1.3	Modelos volumétricos.....	6
1.2.1.4	Modelos híbridos.....	7
1.2.1.5	Modelos de movimiento.....	8
1.2.2	Estrategia de búsqueda.....	8
1.3	Seguimiento no basado en modelos.....	9
1.3.1	Seguimiento basado en regiones.....	9
1.3.2	Seguimiento basado en contornos activos.....	10
1.3.3	Seguimiento basado en rasgos.....	10
1.4	Problemas.....	13
1.4.1	Entrada o salida de Peatones.....	13
1.4.2	Mezcla o separación.....	13
1.4.3	Seguimiento individual.....	13
1.4.4	Seguimiento colectivo.....	13
1.4.5	Oclusión.....	14
2.	Conteo de personas.....	15
2.1	Introducción.....	15
2.2	Conteo basado en modelos de forma.....	16
2.3	Conteo basado en las dimensiones del cuerpo.....	17
2.4	Conteo basado en agrupamiento de trayectorias.....	18
2.5	Conteo basado en la extracción de rasgos.....	18
2.6	Problemas.....	19
3.	Conclusiones.....	19
	Referencias bibliográficas	19
	Anexo 1	22
	Glosario de términos.....	22

Seguimiento y conteo de personas en ambientes exteriores con una cámara fija

Lic. Walter Izquierdo Guerra, Dr. C. Edel García Reyes

Centro de Aplicaciones de Tecnología de Avanzada, 7a #21812 e/ 218 y 222, Siboney, Playa, Habana, Cuba
wizquierdo@cenatav.co.cu

RT_018 CENATAV

Fecha del camera ready: 17 de septiembre del 2009

Resumen: Luego de la detección del movimiento, los sistemas de video-vigilancia están interesados en aplicaciones tales como el seguimiento y conteo de personas. Tales aplicaciones constituyen problemas no resueltos y son de amplio interés investigativo. En los últimos años se ha trabajado muy fuerte en esa dirección, pero las soluciones aún no son tan exactas como se desean. En este reporte proveemos de un estado del arte de los algoritmos más trascendentes en esta rama.

Palabras clave: video-vigilancia, seguimiento, conteo

Abstract: Surveillance systems are interested in some applications such as tracking and counting pedestrians, after the movement detection. Such applications constitute open problems, and scientific community is very interested in it. In the last years, authors have been working in this direction, but solutions are not as good as people expect. In this work, we show a survey of the principal algorithms developed in this branch.

Keyword: Surveillance, Tracking, Pedestrian

1. Seguimiento

1.1 Introducción

Existe un gran número de investigaciones relacionadas con el tema del seguimiento de personas en secuencias de imágenes de video, debido al amplio rango de aplicaciones que estas poseen. Sin embargo, cada aplicación tiene diferentes requerimientos. En particular existe un interés por su aplicación en video vigilancia inteligente.

Luego de la detección del movimiento, los sistemas de video vigilancia generalmente están interesados en darle seguimiento a los objetos en movimiento de un frame a otro. Los algoritmos de seguimiento están estrechamente vinculados con la detección de movimiento durante el procesamiento. El seguimiento pertenece al nivel intermedio de la visión, e involucra el encuentro de relaciones coherentes entre características de la imagen en frames consecutivos con respecto al color, textura, velocidad y posición.

La información del comportamiento de los individuos, para análisis posteriores de interés de video vigilancia, puede ser obtenida de las características de sus trayectorias y de la interacción entre ellos. El análisis de la posición y trayectoria puede determinar ya sea que la persona está en un área prohibida, que está corriendo, saltando u ocultándose. Combinando tal información de dos o más personas, también puede proveer de información de la interacción entre ellas (que las personas estén peleando, una le está apuntando a otra con un arma, etc.).

Los autores de [16] describen técnicas que utilizan la data de la trayectoria para definir un conjunto de actividades que se pueden aplicar a un amplio rango de escenas y ambientes.

El seguimiento robusto de una persona es un problema cuyos enfoques cambian constantemente debido a la complejidad y no rigidez de los movimientos del cuerpo humano, la variación de la iluminación en las imágenes de las secuencias de video, el vestuario, su confusión con el fondo, los problemas de oclusión, etc.

Además, la apariencia de los humanos cambia, no sólo con el cambio del punto de vista sino que está relacionada más fuertemente con las partes visibles del cuerpo humano. Es difícil mantener la identidad de los objetos, cuando los peatones pasan uno muy cerca del otro.

Los algoritmos de seguimiento son en su mayoría dependientes de la aplicación. Se han desarrollado una multitud de herramientas matemáticas para atacar el problema del seguimiento. Algunos de los más prometedores son el filtro de Kalman [19], el algoritmo de condensación [23], el mean shift [28] y técnicas de filtrado bayesiano [22]. El seguimiento puede también envolver vista sencilla o múltiple, cámaras estéreo o monoculares, etc.

Para propósitos de seguimiento en ambientes exteriores, con una sola cámara, los sistemas conocidos que más se adaptan son el LOTS; el cual hace el tracking de blobs aislados de foreground y el W4; que permite el tracking tanto de blobs aislados como de grupos de ellos [25].

La clasificación más relevante de los algoritmos de seguimiento se basa en si estos utilizan o no modelos de forma para seguir los objetos, o sea seguimiento basado en modelos o no.

Por tanto una primera aproximación a dos tendencias principales de seguimiento sería:

1. Seguimiento basado en modelos
2. Seguimiento no basado en modelos

1.2 Seguimiento basado en modelos

La primera clase de método se refiere a todas aquellas técnicas donde el seguimiento de un cierto individuo, en el tiempo se basa en la comparación del contenido de cada imagen con una muestra patrón. Estos algoritmos se enfocan más en el problema de la representación de de la persona, tratando con los cambios de apariencia del mismo. Ellos siguen los objetos por el matching de la proyección de los modelos. Los modelos se construyen off-line con mediciones manuales. Las restricciones de formas se aplican para tratar con la deformación de los blancos y los modelos de movimiento son utilizados para inhibir la ecuación de flujo óptico, resultando en los problemas de optimización en los parámetros del modelo de movimiento.

El seguimiento basado en modelos utiliza generalmente un modelo predeterminado del sujeto para su seguimiento. Se extraen las características de la imagen y se asocian con la estructura del modelo y el movimiento. Sin embargo, esta asociación puede ser una tarea compleja computacionalmente y requiere de una fuerte segmentación del objeto en movimiento desde el background. Tales técnicas son por consiguiente difíciles de implementar en secuencias

borrosas. Pueden ser creados modelos para el seguimiento del cuerpo humano entero y para partes más específicas como manos y rostro.

Mientras los gestos de la mano permanecen medianamente similares sobre la gran variedad de personas (difieren primariamente en el color de la piel), las poses de los cuerpos varían significativamente de una persona a otra y dentro de una misma.

En [2] se describe un método para automáticamente dar seguimiento múltiple, para personas las cuales pueden estar parcialmente ocluidas, paradas o caminando. Su sistema descansa en la detección basada en las partes. Utilizan detectores de cada parte y combinan los resultados de detección de partes para una más robusta detección humana.

La aproximación general para el seguimiento del cuerpo humano basado en modelos utiliza un estilo de predicción-matching-actualización. Primeramente, se predice la pose del modelo en el frame posterior según el conocimiento previo y el historial del tracking. Entonces, el modelo predicho se sintetiza y proyecta en el plano de la imagen para su comparación con la imagen de la data. Para medir la similaridad entre la proyección del modelo y la data, se necesita de una función de evaluación. Según diferentes estrategias de búsqueda, se hace ya sea recursivamente o utilizando técnicas de muestreo hasta que la proyección correcta es finalmente encontrada y se utiliza para actualizar el modelo. La estimación de la pose en el primer frame necesita ser manejada de manera especial. Generalmente, el seguimiento del cuerpo humano basado en modelos involucra tres procesos principales:

- Construcción del **modelo del cuerpo humano**
- Representación del conocimiento previo de las restricciones y **modelos del movimiento**
- Predicción y **estrategia de búsqueda**

1.2.1 Modelos tradicionales utilizados para el cuerpo humano

En la actualidad para la detección y seguimiento de peatones, estos se relacionaron con modelos dinámicos de cuerpos humanos. Estos modelos se adaptan lo más posible al comportamiento de los peatones en casos particulares como son las situaciones de pánico, casos de evacuaciones donde las personas en verdad globalmente se comportan como partículas o fluidos. La ventaja de estos sistemas está fundamentalmente direccionada a su capacidad de reflejar la pose de los peatones en su mayoría.

La construcción de modelos para el cuerpo humano es la base del seguimiento del cuerpo basado en modelos. Generalmente, a mayor complejidad del modelo y mayor precisión de los resultados se observa un mayor gasto computacional. Tradicionalmente, la estructura geométrica del cuerpo se puede representar de las siguientes 4 maneras.

1.2.1.1 Figura esquelética

El seguimiento basado en la figura esquelética se basa en el hecho que la estructura inherente para el movimiento del cuerpo humano es el esqueleto. Por ende la esencia del movimiento humano está típicamente restringida en los movimientos del torso, la cabeza y 4 extremidades, de manera que este método se basa en representar las partes del cuerpo humano como segmentos y unirlos por puntos. Así el movimiento de las piernas, torso, brazos y cabeza pueden ser

aproximado al movimiento del correspondiente segmento de línea. El seguimiento basado en la figura esquelética ha sido desarrollado para el cuerpo humano en su totalidad y para partes de este. Los autores de [8] aproximan los segmentos de líneas en la imagen por la esqueletización del cuerpo entero y calculan el ángulo delimitado por tres componentes principales del cuerpo con la abscisa vertical, el torso, la componente superior de la pierna y la componente baja de esta. En [9] se calcula la transformada de Hough para extraer las líneas que representan las piernas en una secuencia de imágenes. El cambio de inclinación de estas líneas siguen un movimiento armónico simple y este se utiliza como estándar biométrico. Se utiliza el análisis de Fourier para extraer la frecuencia del cambio de inclinación de las piernas y la figura esquelética para calcular el ángulo relativo entre las ellas.



Fig. 1 Figura esquelética

1.2.1.2 Contornos en dos dimensiones (2D)

Los contornos 2D son una representación cerrada de la proyección del sujeto sobre una imagen plana. Las restricciones del movimiento humano, tales como los límites anatómicos de los ángulos de junta, la interpenetración de las partes del cuerpo y posiciones de equilibrio, se le aplican también a las juntas de los contornos 2D, como su contraparte de las figuras esqueléticas.

La principal desventaja de los modelos 2-D es que están restringidos por el ángulo de vista. Para resolver esta desventaja surgen los modelos 3-D (modelos volumétricos). Pero estos requieren muchos más parámetros y por ende un mayor gasto computacional.

[10] y [11] crearon modelos 2D basados en el ajuste de elipses a varios segmentos del cuerpo de la silueta extraída. Entonces derivan vectores característicos tales como la localización del centro de cada elipse, etc.

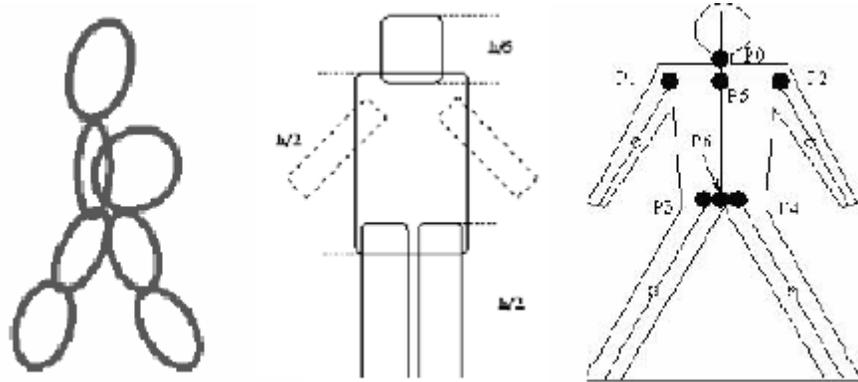


Fig. 2 Contornos 2-D

En [18] construyen un modelo muy simple para la apariencia del cuerpo humano. Este utiliza tanto la información del color como la espacial de los cuerpos. Ellos toman un rectángulo y lo caracterizan por su centro, altura y ancho, haciéndolo corresponder al cuerpo de una persona. Y de las coordenadas espaciales solo actualizan el centro de la caja, de manera que cuando dos rectángulos se unen en uno más grande mantienen invariante su ancho y altura y logran buenos resultados para la oclusión.

1.2.1.3 Modelos volumétricos

Los modelos volumétricos, aunque son computacionalmente más complejos, logran que el seguimiento sea independiente de la vista de la cámara y mejores resultados para este. La representación geométrica del cuerpo humano es más detallada y precisa en estos modelos. Para representar las variadas partes del cuerpo se utilizan cilindros elípticos, conos y esferas. La superficie del cuerpo se representa como una malla poligonal. La ventaja principal de estos modelos es su habilidad para manejar la oclusión. En [12] desarrollaron un modelo para el cuerpo completo con 22 grados de libertad, seis para el dorso y cuatro para cada brazo y pierna. Los sujetos, sin embargo, estaban obligados a usar ropa ceñida al cuerpo con colores contrastados para cada extremidad.

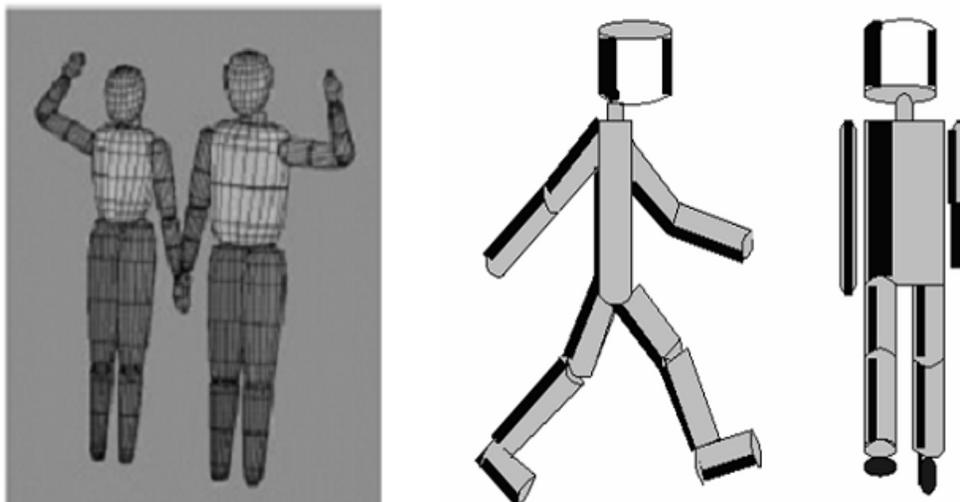


Fig. 3 Modelos volumétricos

1.2.1.4 Modelos híbridos

Muchos modelos híbridos han sido propuestos en la literatura. En [13] se desarrolló un modelo llamado cuerpo clonado, el cual es dinámicamente clasificado respecto al tamaño y textura y emparejado con cada persona para obtener un tracking más robusto tanto de bordes como de textura de las regiones. Este combina modelos 2D y modelos volumétricos. Utilizaron filtrado de partículas (algoritmo de condensación) para calcular los ángulos de junta. Otros autores utilizan el filtro de Kalman aunque este es mucho más complejo.

La mayoría de las técnicas basadas en modelos chocan con el problema de corresponder una imagen humana extraída de una secuencia de frames de video con su representación abstracta por modelos de variada complejidad. Este problema no es trivial, y es gobernado por un número de parámetros de modelo y la eficiencia del algoritmo de segmentación.

En [17] definen un modelo para el cuerpo humano como el mostrado en la fig.4

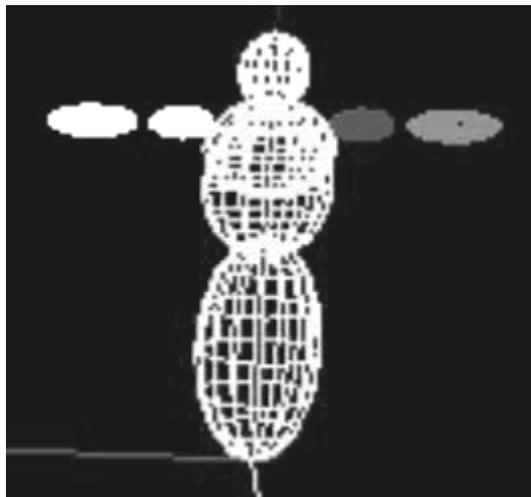


Fig. 4 Modelos híbridos

Los grados de libertad del modelo incluyen la libre rotación y traslación de la cintura, los hombros, los codos y las juntas en general. Cada rotación de junta se representa por los ángulos de Euler, lo que conlleva a un total de 21 grados de libertad y las formas del cuerpo son modeladas como elipses 3-D.

1.2.1.5 Modelos de movimiento

Los modelos de movimiento de las extremidades y sus uniones se utilizan mucho en el tracking. Estos son muy efectivos por las ligaduras que presentan estas extremidades. Estos modelos sirven de conocimiento previo para predecir los parámetros del movimiento o para interpretar y reconocer comportamientos humanos.

1.2.2 Estrategia de búsqueda

La estimación de la pose en un espacio de configuración de alta dimensionalidad es intrínsecamente difícil, de manera que la estrategia de búsqueda tiene que estar acompañada de un diseño cuidadoso para reducir el espacio de solución. Existen cuatro clases principales de estrategia de búsqueda: dinámicas, modelos de Taylor, filtrado de Kalman y muestreo estocástico. Las estrategias dinámicas utilizan fuerzas físicas aplicadas a cada parte rígida del modelo 3-D del objeto seguido. La estrategia basada en los modelos de Taylor mejora incrementalmente una estimación existente, utilizando parámetros del movimiento o diferenciales con respecto a la observación para predecir mejor la búsqueda de direcciones. Este al menos encuentra mínimos locales, pero no garantiza encontrar el mínimo global. Como un estimador lineal recursivo, el filtrado de Kalman puede tratar completamente el seguimiento de la forma y la posición en el tiempo en un caso relativamente libre de desorden en el cuál la densidad de los parámetros del movimiento puede ser modelado satisfactoriamente como una gaussiana. Para manejar la confusión que causa la función densidad de probabilidad por los

parámetros del movimiento ser multimodales y no gaussianos, estrategias de muestreo estocástico, tales como Markov Chain Monte Carlo (MCMC), algoritmos genéticos y Condensación, se diseñan para representar las hipótesis alternativas simultáneas. Entre las estrategias de muestreo estocásticas, la condensación es la más popular.

Comparados con otros algoritmos, los basados en modelos presentan las siguientes ventajas.

- Por la utilización del conocimiento previo de los contornos 3-D o las superficies de los objetos, son intrínsecamente muy robustos. Pueden obtener mejores resultados aún bajo oclusión o interferencia.
- La estructura del cuerpo humano, las restricciones de su movimiento y otros conocimientos previos pueden ser fusionados.
- El tracking basado en modelos 3-D es capaz de extraer la pose del objeto.
- Los algoritmos de tracking basados en modelo 3-D pueden ser aplicados aún cuando los objetos cambien en gran medida las orientaciones durante el movimiento.

Estos presentan desventajas como la complejidad a la hora de la construcción de un modelo y el alto costo computacional.

1.3 Seguimiento no basado en modelos

La idea del tracking no basado en modelos es que la información estructural no siempre es requerida para seguir un objeto y la complejidad puede ser reducida utilizando algún otro método.

1.3.1 Seguimiento basado en regiones

Los algoritmos de seguimiento basado en regiones identifican un blob o región conectada en el espacio que se asocia a cada objeto de interés y se sigue sobre el tiempo utilizando una medida de similitud o un parámetro de correlación cruzada. Para estos algoritmos, el background de la imagen tiene que ser calculado y mantenido dinámicamente y las regiones en movimiento se detectan usualmente por la sustracción de este. Este método sin embargo padece de dos inconvenientes principales. Primeramente, las sombras frecuentemente dan como resultado una representación incorrecta de los blobs. Segundo, en situaciones donde múltiples sujetos están presentes, los indicios de oclusión e individualidad se pierden. Una solución potencial es el uso de varias cámaras.

El algoritmo descrito en [4] utiliza el matching de matrices en dos direcciones (hace un matching de los blobs de el frame actual con los del anterior y viceversa) utilizando el criterio de que cuando dos cajas delimitadas se solapan en un frame, se hace un matching del blob resultante con los anteriores juntos.

Este algoritmo da solución a la evolución de los blobs de un frame a otro analizando los valores del matching de ambos frames, actualiza cada nuevo blob utilizando la información almacenada en los más viejos y mantiene la posición del centroide para formar una trayectoria.

En el método propuesto en [5] se utilizan redes neuronales. Se emplea una red neuronal de retro propagación perceptrón multicapa (RRPM) para el tracking en secuencias de imágenes de baja resolución usando la posición, forma y el color como rasgos. El uso de redes neuronales

provee de un tracking eficiente en situaciones reales particularmente donde la información visual detallada de los peatones no está disponible debido a la baja resolución de la imagen.

Una persona detectada en un frame anterior es buscada en el siguiente usando una red neuronal. Alrededor de la persona detectada se dibujan rectángulos para delimitarlas y se recolectan las variables altura y espesor del rectángulo, conformando así la red neuronal.

Estos algoritmos solamente obtienen los resultados de tracking a nivel de regiones y se encaminan esencialmente hacia la detección del movimiento. En ellos la postura 3-D de los objetos no se adquiere. Consecuentemente, estos algoritmos no pueden satisfacer los requisitos para la vigilancia de un background desordenado o con múltiples objetos en movimiento.

1.3.2 Seguimiento basado en contornos activos

Los algoritmos basados en contornos activos siguen objetos representando sus contornos como bordes bien delimitados y actualizando estos dinámicamente en frames sucesivos. Las serpientes (contornos activos) son funciones suaves que poseen una función de energía interna, definida por su configuración y una configuración de energía externa, definida por la energía de la imagen. Dado un conjunto inicial de puntos de la serpiente, esta intenta lograr una posición que resulte un máximo local de las funciones de energía. Aunque los contornos activos son computacionalmente menos complejos que el seguimiento basado en regiones ellos necesitan ser inicializados como contornos separados de cada sujeto individual. Esta inicialización puede ser bien difícil para objetos complejos. En [14] se usaron serpientes para recubrir los contornos del cuerpo del caminante.

Estos algoritmos logran extraer directamente formas de los sujetos y proveen de una más efectiva descripción de los objetos que los basados en regiones.

En contraste con los algoritmos de seguimiento basado en regiones, los algoritmos de seguimiento basado en contornos activos describen los objetos de una forma más simple, efectiva y reducen el costo computacional. Además bajo escenas complejas u oclusiones parciales, estos algoritmos pueden hacer el seguimiento continuamente. Sin embargo, el seguimiento a nivel de contorno es limitado. La recuperación de la postura 3-D de un objeto, a partir de su contorno en el plano de la imagen es un problema. Otra dificultad es que los algoritmos basados en contornos activos son altamente sensibles a la inicialización del seguimiento, haciéndose muy difícil la inicialización automática.

En [20] utilizan una estructura para acelerar aún más estos algoritmos y su implementación logra algoritmos 15 veces más rápidos que los utilizados hasta el momento.

1.3.3 Seguimiento basado en rasgos

El fundamento del seguimiento basado en rasgos es la reducción de la complejidad funcional cuando se siguen rasgos sobresalientes del objeto, en vez de la región continua del objeto o sus contornos. Por tanto trabajan el reconocimiento y seguimiento de objetos a través de la extracción de elementos, agrupándolos en el más alto nivel de rasgos y luego haciendo un matching entre imágenes. Se utilizan como rasgos parámetros como las esquinas, la información del color y la textura. Existen dos amplias aproximaciones para el tracking basado en rasgos: seguimientos de rasgos dinámicos y estáticos. Se denomina estático cuando los rasgos se extraen

a priori en cada frame y el algoritmo calcula la correspondencia óptima entre ellos. En el seguimiento de rasgos dinámicos los rasgos son determinados y seguidos dinámicamente sobre frames consecutivos, estimando movimiento del rasgo y buscándole en el próximo frame.

Un punto con el seguimiento basado en rasgos es la paradoja entre la complejidad y la eficiencia del seguimiento. Para rasgos de bajos niveles como la posición de las coordenadas de los bordes, que son muy fáciles de extraer, es muy difícil realizar el seguimiento, debido a la dificultad de establecer una correspondencia uno a uno entre ellas. Para rasgos de alto nivel, como blobs y volúmenes 3D, es muy fácil el seguimiento pero muy difícil la extracción.

Estos algoritmos pueden ser subdivididos a su vez en tres categorías según la naturaleza de los rasgos que se seleccionan: algoritmos basados en rasgos globales, basados en rasgos locales, y basados en la dependencia gráfica.

- 1.) Los rasgos que se utilizan en el algoritmo basado en rasgos globales incluye centroides, perímetros, áreas, algunos órdenes de cuadraturas y colores.
- 2.) Los rasgos utilizados en el algoritmo basado en rasgos locales incluyen segmentos de línea, de curvas y vértices de las esquinas.
- 3.) Los rasgos utilizados en los algoritmos basados en la dependencia gráfica incluyen una variedad de distancias y relaciones geométricas entre ellos.

Los tres métodos descritos pueden ser combinados.

En general, como ellos operan en imágenes planas 2-D, los algoritmos basados en rasgos se adaptan satisfactoria y rápidamente para lograr un procesamiento en tiempo real y dar seguimiento múltiple. Sin embargo, los algoritmos basados en la dependencia gráfica no se pueden utilizar en el seguimiento en tiempo real, porque necesitan un tiempo de consumo para la búsqueda y el matching de sus variables. Los algoritmos basados en rasgos pueden manejar la oclusión parcial. Sin embargo, existen varias deficiencias graves de éstos.

- 1.) La tasa de reconocimiento de objetos basada en los rasgos de las imágenes en 2-D es baja, debido a la distorsión por la no linealidad de la perspectiva de la proyección.
- 2.) Estos algoritmos son generalmente incapaces de recobrar la postura 3-D de los objetos.
- 3.) No tiene un tratamiento estable para la oclusión, solapamiento e interferencia de estructuras.

En [3] se usa un algoritmo de segmentación del background en combinación con un clasificador de objetos para rápidamente encontrar peatones en cada frame del video. Luego de que el posible peatón es detectado, el objeto en movimiento es subdividido en tres zonas (cabeza, parte alta del cuerpo, parte baja del cuerpo) y se extraen el color y las propiedades espaciales de cada parte las cuales forman el modelo de apariencia básico en este sistema. La información del color se analiza en el espacio de colores HSI (matiz, saturación, intensidad en inglés).

La efectividad del tracking depende fuertemente de la selección de los rasgos. Su clasificador utiliza el detector para obtener la información básica espacial de un posible objeto y utiliza el modelo de background para adicionar mayor información. El detector provee de un

conjunto de ventanas de detección en un frame. Cada ventana está completamente definida por su dimensión y posición. Primeramente dividen cada ventana de detección en tres zonas del cuerpo individuales (cabeza, parte alta del cuerpo, parte baja del cuerpo) utilizando la razón (1/4, 3/8, 3/8 respectivamente). Luego cada parte es procesada hasta que su contenido principal consista en una región conectada de píxeles del foreground. Luego se extrae la información del color. El vector característico de cada parte del cuerpo contiene la posición, la dimensión y el histograma.

Un enfoque muy interesante es el publicado por [1]. En él se combinan los métodos de procesamiento de imágenes con los modelos dinámicos del comportamiento de los peatones. La meta es detectar peatones prestando atención a su dinámica y comportamiento. De manera que necesitan información preferentemente sobre sus desplazamientos que sobre su apariencia. Teniendo la posición hipotética de un peatón y la correspondiente región de la imagen en un frame, se computa su correlación con la región correspondiente en el frame que le sigue. El máximo de correlación da la localización de el mejor matching entre las dos regiones. El vector identificado por la posición corresponde al vector desplazamiento de la región de la imagen actual sobre los dos frames. Lo interesante entre este bien conocido método es que en dos frames consecutivos un ser humano puede cubrir una distancia limitada, de manera que es razonable pensar que la región de búsqueda, usada para computar la correlación, contiene la verdadera posición del blanco. Se aplica entonces restricciones de comportamiento generadas por los vectores de movimiento.

Para cada par de frames consecutivos computan el vector desplazamiento por la maximización de la correlación. Suponiendo que la altura promedio de una persona es 1.70m, se obtiene un ajuste automático del tamaño de la región del blanco hipotético en la imagen. Esta simple maña evita definir modelos de deformación más complejos.

En [19] utilizan también el seguimiento basado en rasgos. Para seguir el movimiento de estos, utilizan el filtro de Kalman y lo combinan con técnicas de optimización para la asociación de la data. La correspondencia entre la predicción de cada rasgo en una nueva medición la basan en la distancia de Mahalanobis.

Hasta aquí el esquema taxonómico quedaría de la manera que se muestra en la fig. 5:

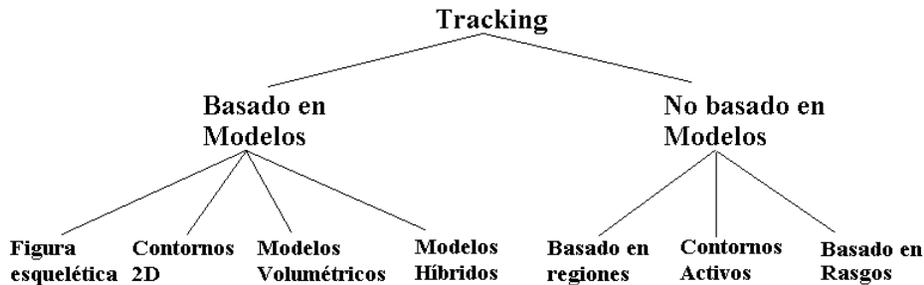


Fig. 5 Esquema taxonómico del seguimiento de peatones

Con ello no se quiere decir que la utilización de uno de estos algoritmos excluya el uso de otro. De hecho existen una cantidad innumerable que utilizan la combinación de ellos para lograr

una mayor robustez en el tracking. Siempre teniéndose en cuenta el gasto computacional que esto implica.

En [26] utilizan tres algoritmos cooperantes. Uno basado en contornos activos, otro basado en regiones y otro basado en la detección de la cabeza (contorno 2D) para una mejor inicialización de las trayectorias.

Todas estas metodologías se enfrentan a su vez con tres problemáticas esenciales: la **entrada o salida de peatones** a la escena, la **mezcla o separación**; ya sea de personas individuales como de blobs en general y uno estrechamente enlazado que es la **oclusión**.

1.4 Problemas

1.4.1 Entrada o salida de Peatones

La solución a este problema parte de la premisa de que durante las oclusiones, los blobs individuales sumergidos dentro del grupo, se suponen siempre forman parte del grupo. Por lo tanto, los objetos no pueden desaparecer de la escena a menos que lo hagan a través de los bordes predefinidos. Estos bordes están definidos como zonas de la imagen a través de las cuales los objetos pueden llegar o abandonar la escena. De esta manera se pueden detectar personas que aparecen o se ocultan, así como objetos recogidos y abandonados [4].

1.4.2 Mezcla o separación

El tracking de personas enfrenta la problemática de la mezcla o separación de blobs. Existen dos tendencias principales a trabajar en esta dirección: el seguimiento individual o colectivo. Estas dos modelaciones marcan tendencias actuales del tratamiento del problema.

1.4.3 Seguimiento individual

Los algoritmos se centran en una persona o blob, y aún cuando estos son sumergidos en otro, mantienen sus propiedades y coordenadas almacenadas, de manera que aún sumergidas en un grupo, el tracking sigue siendo para cada blob por separado.

En [5], cuando una multitud de personas son en parte ocluidas por ellas mismas, la imagen del foreground del grupo de personas se divide en individuales utilizando una RRPM. Esta red incorpora tres técnicas diferentes que descansan en una línea que conecta los píxeles superiores de la imagen binaria de foreground, la proyección vertical de la imagen binaria del foreground y el valor de la varianza de los píxeles en las áreas divididas.

1.4.4 Seguimiento colectivo

Cuando un blob se sumerge dentro de otro, el algoritmo olvida las coordenadas individuales y prosigue la trayectoria con las propiedades colectivas del nuevo grupo.

El algoritmo propuesto en [4] trabaja con la posición relativa de los blobs y sus interacciones en la formación o disolución de grupos y no mantiene la información de la posición de los blobs cuando forman un grupo.

Si dos blobs se mezclan para formar uno nuevo, este blob particular es clasificado como un grupo. Este nuevo grupo es seguido individualmente, aunque la información de los dos blobs mezclados es almacenada para usarle posteriormente. Si el grupo se separa nuevamente, el sistema usa la dirección de la velocidad y las características del blob para identificar correctamente ambos blobs por separado. Siempre es posible una interpolación de la posición del blob seguido en aquellos frames donde formaba parte de un grupo para obtener las trayectorias completas.

Por ende, cuando dos blobs se unen formando un grupo, la información se mantiene pero el blob que se sigue es el del grupo. Esto significa que la información acerca del centroide de los blobs individuales no está disponible mientras ellos formen parte del grupo.

1.4.5 Oclusión

El fenómeno de oclusión se ha convertido en un objetivo esencial de investigación en nuestros tiempos. Muchos trabajos recientes presentan tratamientos particulares para este fenómeno [2, 4, 5, 21, 24, 27]. A pesar de los avances existentes y del alcance de muy buenos resultados en la comunidad científica, no existe una solución global para este problema.

En [27] utilizan un algoritmo de seguimiento basado en regiones, y lo enfocan directamente a la oclusión. Ellos logran buenos resultados antes, durante y después del fenómeno.

[4] trata las oclusiones completando las trayectorias mediante una interpolación de la posición de los centroides. Cuando la oclusión se produce por un objeto fijo, se utiliza la velocidad media de los frames anteriores. Si la oclusión es debida a una unión de blobs, entra a considerarse entonces el centroide del nuevo blob.

Se predice la posición, usando los valores almacenados de posición y velocidad, para completar la trayectoria de un blob mientras se agrupa con otros. De esta manera, una lista consistente temporal de los blobs se mantiene junto con su trayectoria y sus posiciones.

El algoritmo propuesto en [2] no espera encontrar todas las partes del cuerpo humano, su sistema explícitamente trabaja sobre la oclusión de partes por la consideración de la detección de todos los objetos juntos, de manera que es capaz de detectar una persona por la detección individual de alguna de sus partes.

Los peatones son seguidos basándose en la detección completa dónde sea posible. En presencia de estas oclusiones, sólo algunas partes se pueden ver; en tales casos, el sistema sigue las partes visibles y combina los resultados de asociaciones de partes para el seguimiento humano. Cuando no hay disponible una detección confiable, se aplica el tracker mean shift [28]. Para una oclusión completa por otros humanos u objetos de la escena, el tracking se infiere de las observaciones antes y después de tal oclusión. Este método no requiere inicialización manual; en lugar de eso las trayectorias son inicializadas y terminadas automáticamente basadas en las salidas de detección.

[5] Trata la oclusión dividiendo la mancha de píxeles de un grupo de personas en regiones individuales basándose en tres metodologías principales:

- 1) Solapamiento de líneas:

Se asume que las personas tienen una postura aproximadamente recta. En este caso la división puede ser limitada solo a la dirección vertical. La línea de solapamiento se dibuja conectando los píxeles superiores de las columnas de foreground. Luego se toma el mínimo local y se separa verticalmente. Para ello se hace necesario que la línea de solapamiento pueda ser suavizada para librarle de ruido.

2) Proyección vertical

Es la suma de todas las filas de la imagen binaria. Se construye el perfil y se le aplica la metodología de la línea de solapamiento.

3) Varianzas de regiones divididas

Se basa en el principio de que si para un objeto existe una varianza calculada a priori, todo grupo de píxeles o región que se le adicione a este objeto, la incrementará sustancialmente de su valor original.

Como los tres métodos discutidos chequean diferentes aspectos de la imagen, se integran en una red neuronal para obtener la mejor división.

2. Conteo de personas

2.1 Introducción

En los últimos años se han iniciado una gran cantidad de investigaciones y trabajos en el área del procesamiento de imágenes con el objetivo de obtener mayor precisión y confiabilidad en los procesos de estimación del conteo de la cantidad de personas.

A pesar de la cantidad de métodos presentes en la literatura, el problema está lejos de ser resuelto y la mayoría de las dificultades conocidas radica en el paso primario. No existe un método robusto de detección de objetos.

El conteo de personas es una importante tarea en los sistemas automáticos de video vigilancia. Puede ser una sub-tarea en un estado determinado de un procesamiento, o puede ser una tarea de interés primario. La estimación robusta de la cantidad de personas puede proveer de respuestas a preguntas tales como: ¿cuántas personas estuvieron dentro de una habitación entre dos tiempos determinados?

Existen muy diversas maneras de tratar el problema del conteo de personas. Ellas se dividen en 4 metodologías principales:

- 1 Conteo basado en modelos de forma
- 2 Conteo basado en las dimensiones del cuerpo (Área o cantidad de píxeles)
- 3 Conteo basado en agrupamiento de trayectorias
- 4 Conteo basado en la extracción de rasgos

Aunque genéricamente se encuentran separados, existe una relación muy estrecha entre todos estos procedimientos y es muy común encontrar trabajos donde combinan varios de ellos. De hecho, en los basados en la extracción de rasgos nos encontramos con la posibilidad de que el área o la densidad de píxeles sea extraída como una característica más.

Existe además una fusión más general. La idea es que el agrupamiento de trayectorias, a pesar de ser una rama independiente, pues no necesita saber cómo fueron estas obtenidas, puede ser el resultado de un seguimiento basado en rasgos, modelos, etc. Todo está basado en el mecanismo de detección que utilicemos. Es por ello que aunque en una buena aproximación, existen diferentes tendencias para tratar el problema, en el nivel de conocimiento que está esta ciencia, no existen caminos aislados que obtengan un robusto resultado, sino una combinación y balance de metodologías y rendimiento.

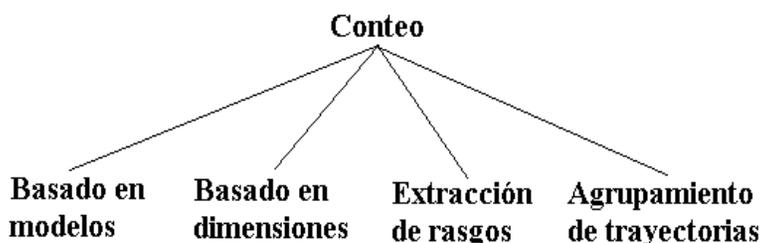


Fig.6 Esquema Taxonómico de conteo de personas

2.2 Conteo basado en modelos de forma

Se basan fundamentalmente en la detección de formas para la identificación previa de partes humanas y a partir de ello utilizarles como método de contabilidad de individuos. Una de las más utilizadas es la semejanza existente entre la parte de los hombros y la cabeza con la letra “Ω”. Luego de hacerse la correspondiente identificación, pues simplemente contando las cabezas, se obtiene el número de individuos presentes en la escena.

En [33] toman en cuenta esta aproximación. Utilizan un método rápido de búsqueda de contornos para detectar los posibles candidatos a la forma de la letra “Ω”. Se aplica un filtro detector de bordes de Canny y se representa el contorno como un conjunto de 23 puntos.

Otra manera de utilizar esta metodología es el método expuesto en [34]. Proponen una aproximación muy interesante, pues después de la detección presenta un algoritmo de reducción de ruido (eliminar objetos detectados que no son personas).

Su modelo se basa en el hecho que cada humano por separado (la mayor parte del tiempo) tiene una abscisa mayor en la dirección vertical. Por tanto definen una persona como un objeto con un ángulo con respecto a la vertical y una excentricidad media. Definen los intervalos de valores alto medio y bajo para el ángulo, la excentricidad y área que ocupa. A partir de estas características construyen una tabla de clasificación, donde por ejemplo; un objeto que entre en la categoría baja para estas tres características es considerado ruido, pues una persona no puede ser más pequeña que el umbral que ellos definen, con una postura casi horizontal y con muy baja

excentricidad al mismo tiempo. Luego con esta filosofía rellenan la tabla con los valores de probabilidad de que sea persona o ruido y así obtienen su mecanismo de clasificación.

Para el conteo, utilizan los features extraídos anteriormente para cuantificar la cantidad de personas presentes en un blob. Por ejemplo, la excentricidad es umbralada para una persona, y por una extrapolación del cálculo de esta se estima el número de individuos. Además definen otras variables como el factor de calidad, construyen otras tablas de clasificación y obtienen muy buenos resultados en su algoritmo.

La principal desventaja es la necesidad de preparar manualmente el conjunto de aprendizaje. Sería interesante darle un tratamiento más automatizado, donde el sistema sea capaz de entrenarse y aprender por sí mismo.

2.3 Conteo basado en las dimensiones del cuerpo

El conteo basado en las dimensiones del cuerpo depende de la densidad de píxeles del foreground. Normalmente requieren de una imagen de referencia donde no haya personas presentes. Existen variantes en las cuales se les provee de imágenes con variadas cantidades de personas para entrenar el algoritmo utilizado.

Un enfoque muy interesante es el expuesto en [29] donde se extraen los píxeles de foreground utilizando el filtro de la mediana. Luego aplican operaciones morfológicas para suavizar el resultado. Para cada cámara, crean tres rejillas de tamaño diferente y las aplican a las imágenes binarias para obtener tres conjuntos de datos. Cada imagen capturada se va a almacenar con una etiqueta (el número estimado de personas) y una cantidad de valores igual al número de celdas que componen las rejillas que van a almacenar la cantidad de píxeles del foreground de cada celda.

Implementan los siguientes algoritmos: (a) regresión lineal múltiple; (b) vecino k más cercano con la distancia euclidiana; (c) una red neuronal retroalimentada de dos capas; (d) una red neuronal probabilística de dos capas; (e) máquina de soporte vectorial. Buscan un clasificador para determinar si no hay ninguna persona en una imagen y uno para determinar la cantidad de personas en una imagen. Entrenan el algoritmo con imágenes que contienen diferente cantidad de personas y buscan resultados.

En [32] proponen dos estrategias para el conteo. La primera utiliza dos umbrales, el ancho promedio de los blobs, que constituyen a una sola persona y el área promedio de la región superior de los blobs correspondiente a la zona de la cabeza, los cuales son el resultado del entrenamiento previo. En la segunda, se divide la región de la cabeza en 10 regiones verticales de igual tamaño. Para cada una se calcula el número de píxeles de foreground y se divide por el área total de la subregión. Los rasgos extraídos de la zona, más el ancho de la región de la cabeza conforman un vector de 11 dimensiones. Luego, a los vectores característicos de los objetos en movimiento se les hace un matching contra vectores de referencia almacenados con un clasificador no paramétrico. Este método va acompañado de un entrenamiento previo. Este método es muy eficiente en cuanto a gasto computacional.

[36] presenta dos algoritmos de diferente naturaleza, uno de ellos entra en esta clasificación. En este, calculan el valor medio del área que ocupa una persona en una imagen en un número suficientemente grande de muestras para entrenar el procedimiento. Luego cuando tienen un blob que constituye un grupo de personas, estiman el número de estas a través de la división del área total del blob por el valor medio estimado.

2.4 Conteo basado en agrupamiento de trayectorias

La principal característica de esta metodología es su basamento en el agrupamiento de trayectorias. En su generalidad los trabajos relacionados asumen que dos trayectorias muy similares solo pueden corresponder a partes diferentes de un mismo individuo o lo que es igual que las trayectorias extraídas de un mismo objeto, son más similares unas con otras que las extraídas de otro cualquiera. El problema está entonces en establecer cuando son suficientemente similares. Esta es la idea principal del tratamiento que le dan al problema en [37].

En [30] el proceso de tracking se trata como una caja negra que lo que hace es proveernos de la data. Aplican un algoritmo de clusterización jerárquica para reducir la diferencia entre el número de trayectorias y el número de personas presentes en la escena, para lo cual se propone la utilización de técnicas de agrupamiento jerárquico. Se consideran dos representaciones diferentes de trayectorias (ICA {Independent component analysis, english} y máxima correlación cruzada) y diferentes procedimientos de agrupamiento para cada una de ellas, basados en dos medidas de distancias diferentes (Harsdorff y euclidiana respectivamente).

2.5 Conteo basado en la extracción de rasgos

Estos algoritmos se basan en la extracción de ciertas características que identifican a un individuo. Los rasgos más comunes son los bordes, la información del color y la textura.

En [31] segmentan las componentes que presentan un movimiento homogéneo. Luego extraen un conjunto de características de cada región segmentada y la correspondencia entre estas y el número de personas por segmentación se entrena con un proceso de regresión gaussiana.

Los rasgos extraídos los dividen en tres tipos fundamentales:

- Rasgos de segmentación
- Rasgos de borde interno
- Rasgos de textura

Los de segmentación guardan la información de la forma y tamaño de la región segmentada, ya sea el área, perímetro, orientación de los bordes, así como la razón entre el perímetro y el radio.

Los de borde interno, en áreas densas, contienen la información del número de personas en el área segmentada. Entre ellos están el número total de píxeles de borde, la orientación de estos, entre otros.

En los de textura tratan con funciones de homogeneidad, energía y entropía.

El método propuesto trae una muy fuerte dependencia de la segmentación y por ende de una buena detección. Aunque presentan muy buenos resultados, el problema de la detección no está aún robustamente resuelto y es por ello que constituye una buena aproximación pero no una solución universal al problema.

2.6 Problemas

Los problemas encontrados en el conteo de peatones están muy solapados con los expuestos para tracking. Esta afirmación parte del hecho de la similitud que existe entre los principales algoritmos de tracking y conteo, ya que lo que se busca es particularizar a una persona en un blob de píxeles detectados como foreground. Sea cual sea la aplicación de estas que deseemos, vamos a tener una fuerte dependencia de esta particularidad. Por ejemplo si tenemos una oclusión, esto continuará constituyendo problema para el conteo, pues el simple hecho de que no estén bien definidos los límites de cada peatón hace que el modelo que se le aplique falle, al igual que lo hacía para el tracking. Es por ello que no resulta difícil extrapolar los problemas de tracking a conteo, basándonos en la heurística de que es necesaria la caracterización de un peatón sea cual sea el procedimiento que hayamos escogido para ello.

3. Conclusiones

Tanto el tracking de peatones como el conteo de los mismos, han pasado por muchas etapas de investigación. La diversidad de herramientas para enfrentar estos problemas ha sido colosal, pero la realidad es que siguen constituyendo problemas abiertos para la comunidad científica. Como se ha visto en este trabajo, ya se han definido una serie de problemas particulares a resolver, que aun siguen siendo de mucho interés. Las aproximaciones siguen siendo cada día más exactas pero no tenemos la solución global en nuestras manos.

La mayor parte de los trabajos concluyen en el hecho de que no existe una detección robusta para los peatones y es por ello que su sistema no da los mejores resultados. Esta afirmación no es del todo errada, pues si no existe una buena segmentación de la persona, la mayoría de los modelos fallan. En fin en los sistemas de video vigilancia aún queda mucho por hacer, en especial el conteo y tracking de peatones dependen de etapas anteriores de procesamiento y es por ello que se deben ver no como problemas aislados sino como un algoritmo general. Es necesario establecer estrategias de modelación, segmentación, tracking y conteo.

Referencias bibliográficas

1. Antonini, G. y Venegas, S. 2006. Priors for Detection and Tracking of Pedestrians in Video Sequences. *International Journal of Computer Vision*.
2. Wu, B. y Nevatia, R. 2007. Detection and Tracking of Multiple, Partially Occluded Humans by Bayesian Combination of Edgelet based Part Detectors. . *International Journal of Computer Vision*.
3. Seitner, F.H. y Hanbury, A. 2006. Fast pedestrian tracking based on spatial features and colour. *Pattern Recognition and Image Processing Group, Vienna University of Technology, Vienna, Austria*.
4. Fuentes, L. M. y Velastin S. A. 2006 People tracking in surveillance applications
5. Do, Y. 2005. Tracking People in Video Camera Images Using Neural Networks
6. Masoud, O. y Papanikolopoulos, N. P. 2001. A Novel Method for Tracking and Counting Pedestrians in Real-Time Using a Single Camera. *IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY*.

7. Singh, M. 2004. Robust Tracking and Human Activity Recognition. Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Canada.
8. Ali, A. y Aggarwal J. K., 2000. Segmentation and Recognition of Continuous Human Activity. Computer and Vision Research Center.
9. Cunado, D. et al., 1997. Using Gait as a Biometric, via Phase-Weighted Magnitude Spectra. Department of Electronics and Computer Science, University of Southampton, Highfield, Southampton SO17 1BJ, England.
10. Lee, L., 2001. Gait Analysis for Recognition and Classification. Artificial Intelligence Laboratory, Massachusetts Institute of Technology.
11. Shakhnarovich, G., 2001. Integrated Face and Gait Recognition From Multiple Views. Artificial Intelligence Laboratory, Massachusetts Institute of Technology.
12. Gavrilu, D. M. et al. 1996. 3-D model-based tracking of human in action: a multi-view approach. Computer Vision Laboratory, University of Maryland.
13. Green, R. D., 2006. Quantifying and Recognizing Human Movement Patterns from Monocular Video Images - Part I: A New Framework for Modeling Human Motion. Human Interface Technology Lab, University of Canterbury, Christchurch, New Zealand.
14. Niyogi, S. A. y Adelson, E. H., 1994. Analyzing gait with spatiotemporal surfaces.
15. Hu, W. y Tan, T., 2004. A Survey on Visual Surveillance of Object Motion and Behaviors. IEEE.
16. Morris, B. T. y Trivedi, M. M. 2008. A Survey of Vision-Based Trajectory Learning and Analysis for Surveillance. IEEE
17. Chu, C. W. y Nevatia, R. 2008. Real-Time 3D Body Pose Tracking from Multiple 2D Images. Institute for Robotics and Intelligent System, University of Southern California.
18. Wang, H. y Suter, D. 2007. A consensus-based method for tracking: Modelling background scenario and foreground appearance. Department of Electrical and Computer Systems Engineering, Monash University, Clayton Vic. 3800, Australia
19. Pinho, R. R. et al. 2007. An Efficient and Robust Tracking System using Kalman Filter. Faculdade de Engenharia da Universidade do Porto, Portugal.
20. Dou, Y. y Xu, J. 2007. FPGA-Accelerated Active Shape Model for Real-Time People Tracking. Department of Computer Science, National University of Defence Technology, Changsha, P.R. China.
21. Wu, B. y Nevatia, R. 2006. Tracking of Multiple, Partially Occluded Humans based on Static Body Part Detection. University of Southern California, Institute for Robotics and Intelligent Systems, Los Angeles.
22. Rius, I. et al. 2005. Action Spaces for Efficient Bayesian Tracking of Human Motion. Centre de Visió per Computador, Edifici O. Campus UAB. 08193, Bellaterra, Spain.
23. Rius, I. et al. 2005. A 3D Dynamic Model of Human Actions for Probabilistic Image Tracking. Centre de Visió per Computador/Department of Computer Science, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain.
24. Do, Y. 2005. Region Based Detection of Occluded People for the Tracking in Video Image Sequences. School of Electronic Engineering, Daegu University, Kyungsan-City, Kyungpook, 712-714, Korea.
25. Haritaoglu, I et al. 2000. W4: Real-Time Surveillance of People and Their Activities. IEEE.
26. Siebel, N. T. y Maybank, S. 2001. Fusion of Multiple Tracking Algorithms for Robust People Tracking. Computational Vision Group, Department of Computer Science, The University of Reading, Reading RG6 6AY, England.
27. Sindhu, A. J. y Morris, T. 2004. A Region Based Approach To Tracking People Before, During And After Occlusions, Department of Computation, UMIST, Manchester, U.K.
28. Velipasalar, S. et al. 2006. Automatic Counting of Interacting People by Using a Single Uncalibrated Camera. Princeton University, Electrical Engineering Dept.

29. Roqueiro, D. y Petrushin, V. A. 2006. Counting people using video cameras. The International Journal of Parallel, Emergent and Distributed Systems.
30. Biliotti, D. ; Antonini, G. y Thiran, J. P. 2005. Multi-layer hierarchical clustering of pedestrian trajectories for automatic counting of people in video sequences.
31. Antoni, B. et al. 2008. Privacy Preserving Crowd Monitoring: Counting People without People Models or Tracking. IEEE Conf. on Computer Vision and Pattern Recognition, Anchorage.
32. Valle, J. D. et al. 2007. People Counting in Low Density Video Sequences. Springer-Verlag Berlin Heidelberg.
33. Sidla, O. et al. 2006. Pedestrian Detection and Tracking for Counting Applications in Crowded Situations. Proceedings of the IEEE International Conference on Video and Signal Based Surveillance (AVSS'06)
34. Panagiotakis, C. et al. 2006. Automatic People Detection and Counting for Athletic Videos Classification. Computer Science Department, University of Crete, Heraklion, Greece.
35. Kong, D. et al. 2006. A Viewpoint Invariant Approach for Crowd Counting. IEEE.
36. Kilambi, P. et al. 2008. Estimating pedestrian counts in groups. Department of Computer Science and Engineering University of Minnesota, Twin Cities, USA.
37. Rabaud, V. y Belongie, S. 2006. Counting Crowded Moving Objects. Department of Computer Science and Engineering, University of California, San Diego.

Anexo 1

Glosario de términos

Provenientes del inglés:

Blob: agrupación de píxeles detectados como objeto y que pueden constituir uno o varios individuos

Frame: conocido como “marco” en español, es una imagen específica de una secuencia de video

Background: todo aquello de una escena que no constituye un objeto de interés

Foreground: píxeles detectados como objetos de interés en una escena

Tracking: seguimiento de un objeto de interés en una secuencia de video

Matching: comparación entre imágenes o partes de ellas

Features: rasgos extraídos de las imágenes que caracterizan el objeto de interés

RT_018, diciembre 2009

Aprobado por el Consejo Científico CENATAV

Derechos Reservados © CENATAV 2009

Editor: Lic. Lucía González Bayona

Diseño de Portada: DCG Matilde Galindo Sánchez

RNPS No. 2142

ISSN 2072-6287

Indicaciones para los Autores:

Seguir la plantilla que aparece en www.cenatav.co.cu

C E N A T A V

7ma. No. 21812 e/218 y 222, Rpto. Siboney, Playa;

Ciudad de La Habana. Cuba. C.P. 12200

Impreso en Cuba

