

## USO DE FLUJO ÓPTICO PARA LA DESCRIPCIÓN DEL ENTORNO DINÁMICO DE UN VEHÍCULO URBANO AUTÓNOMO

Genaro Ellian Herrera Bandin, [ellianhbandin@gmail.com](mailto:ellianhbandin@gmail.com)<sup>1 2</sup>

João Pedro Finoto, [jpfm2210@gmail.com](mailto:jpfm2210@gmail.com)<sup>1</sup>

Ingrid Lorena Argote Pedraza, [ingridargote@usp.br](mailto:ingridargote@usp.br)<sup>1</sup>

Prof. Marcelo Becker, [becker@sc.usp.br](mailto:becker@sc.usp.br)<sup>1</sup>

<sup>1</sup>Escola de Engenharia de São Carlos - Universidade de São Paulo, Av. Trabalhador São-carlense, 400, Pq Arnold Schimidt São Carlos - SP/Brasil, CEP 13566-590

<sup>2</sup>Universidad Autónoma de Ciudad Juárez, Av. Plutarco Elías Calles #1210, Fovissste Chamizal, Ciudad Juárez, Chihuahua, México, C.P. 32310

**Resumo:** *El siguiente trabajo presenta un algoritmo que describe el entorno dinámico de un vehículo urbano autónomo interpretando los datos del flujo óptico de la escena. Las pruebas se realizaron dentro del proyecto SENA (Sistema Embebido de Navegación Autónoma) de la Escola de Engenharia de São Carlos de la Universidade de São Paulo. El sistema cuenta con una cámara monocular a color con una resolución de 640x480 que funciona a una velocidad de 25 fps. El algoritmo identifica las esquinas utilizando el método de Shi-Tomasi. Estos datos son almacenados para compararlos con los datos de la siguiente imagen y posteriormente obtener el flujo óptico con el método de Lucas-Kanade. La orientación y la longitud de los vectores de flujo óptico son discriminados para determinar si representan objetos estáticos u objetos dinámicos. Los objetos dinámicos son segmentados para posteriormente fusionarse con la segmentación del sensor láser del vehículo (Velodyne). Los vectores de flujo óptico de los objetos estáticos son generados por el movimiento del vehículo, por lo que se puede obtener su velocidad y su orientación con esta información. El algoritmo determina el foco de expansión (FoE) del flujo óptico con el cual se puede ubicar el horizonte en la imagen. Obteniendo la posición del FoE, se determina un factor de proporcionalidad con el cual la magnitud de los vectores de flujo se normaliza, ya que los vectores más alejados del foco de expansión aparecerán más alargados. Este algoritmo permite la localización de la línea del horizonte. Este dato se requiere para describir mejor un ambiente externo por el cual vehículo tiene que transitar, por ejemplo en calles con diferentes grados de inclinación. Los resultados de este sistema son fusionados con los resultados de otros algoritmos para aumentar la confiabilidad de los mismos.*

**Palabras clave:** *Vehículos autónomos, Sistemas de visión, Flujo óptico, Odometría Visual*

### 1. INTRODUCCIÓN

Cada año mueren 1.24 millones de personas a nivel mundial debido a accidentes de tránsito y entre 20 y 50 millones quedan incapacitadas debido a estos percances como lo muestran las estadísticas de la WHO (2013). La mitad de esos fallecimientos son peatones, ciclistas y motociclistas. La gran mayoría de estos incidentes se debe a la falta de habilidad del conductor al manejar. Los desarrollos realizados en áreas como la robótica móvil, la electrónica y la instrumentación tienen una gran participación en el sector automotriz que posteriormente forman parte de la vida cotidiana proporcionando servicios de calidad que garantizan la seguridad del usuario. Estos avances permitirán la automatización total o parcial de las diferentes funciones al momento de conducir el vehículo, aprovechando la velocidad y confiabilidad de los sistemas robóticos. Ejemplos de estas funciones incluyen algoritmos para mantener al vehículo dentro de un carril, guardar distancia entre los vehículos, mantener las condiciones de velocidad adecuadas y respecto a las normas. Conducir a través de las vías, sobrepasar otros vehículos, evitar obstáculos entre otras.

Actualmente, los progresos realizados en el desarrollo de vehículos autónomos les permite ser capaces de transitar y tomar decisiones. Por ejemplo, el vehículo autónomo desarrollado por Google el cual utiliza un software de inteligencia artificial que puede detectar cualquier objeto que se encuentre cerca al vehículo; este proyecto ha conducido más de 1000 millas sin ningún tipo de intervención humana. Dicho software de inteligencia artificial es alimentado por un sensor LIDAR, Cámaras de video, estimadores de posición, radares entre otros sensores (IEEE Spectrum 2010). Al igual que Google, existen muchas otras compañías interesadas en investigar y desarrollar este tipo de tecnologías. Tal es el caso del Chevrolet Tahoe SUV llamado Boss, desarrollado por el equipo Tartan Racing de la universidad de Carnegie Mellon y la corporación General Motors para participar en el DARPA Urban Challenge, el cual fue el ganador cumpliendo un

recorrido de 90 km sin ocasionar mayor accidente. Form (2008) menciona que ese mismo año en la Universidad Técnica de Braunschweig, fue presentado un vehículo autónomo llamado "Caroline", desarrollado por el equipo CarOLO, el cual realiza complejas maniobras de conducción autónoma. Este tipo de proyectos han demostrado su funcionalidad tanto dentro de ambientes urbanos como en terrenos *off-road* como se ha demostrado en las competiciones DARPA Grand Challenge y DARPA Urban Challenge. Los detalles pueden ser encontrados en el sitio web de DARPA (2014) y en las publicaciones de Buelher M. Sanjiv y S. Iagnemma (2005 y 2009). El desarrollo de tecnologías enfocadas a la navegación autónoma y su implementación en el sector automotriz comercial puede ayudar en un futuro cercano a disminuir de manera importante la cifra de accidentes a nivel mundial.

Antes de pensar en realizar la navegación autónoma, es indispensable lograr determinar el desplazamiento del robot móvil y su ubicación a través del tiempo. Para lograrlo, se requiere de un método capaz de realizar la odometría. Actualmente existen varios métodos para esto, como el uso de encoders en las ruedas o la localización satelital. Sin embargo, ambos métodos presentan varios inconvenientes que comprometen la calidad de sus resultados, por ejemplo, el resultado de los encoders puede verse afectado debido al derrape de las ruedas o a problemas mecánicos. Niester *et al.* (2004) sostiene que los sistemas de localización satelital presentan desventajas para realizar la odometría ya que el correcto funcionamiento de la tecnología GPS depende de las condiciones ambientales, no logran hacer una correcta triangulación dentro de cañones urbanos, no funcionan adecuadamente en aplicaciones en ambientes cerrados tales como túneles o dentro del laboratorio además de que su uso puede tener un costo adicional. Debido a estos inconvenientes, es necesario buscar un método capaz de realizar la odometría que sea de bajo costo y que pueda superar las limitaciones que presentan otros algoritmos y sensores utilizados con el mismo fin. Es importante que el método que realiza la odometría tenga una baja demanda computacional para optimizar los recursos y que estos puedan ser aprovechados para las diferentes tareas que realiza el vehículo autónomo pero sin menguar la calidad de los datos obtenidos. Por este motivo se plantea en el presente trabajo describir el entorno dinámico del vehículo con la ayuda de un sensor visual monocular.

Una de las áreas de oportunidad de los vehículos autónomos y en la robótica móvil en general, es la implementación de los sistemas de visión para el sensado del entorno. Los sensores de visión representan una alternativa de bajo costo comparados con sensores que cumplen con funciones similares tales como sonares, acelerómetros, encoders, dispositivos de localización satelital, y dispositivos que implican tecnologías láser. También, como se muestra en el trabajo de Broggi *et al.* (2013), los sensores de visión presentan ventajas tales como un menor impacto en el diseño del vehículo y en ocasiones obteniendo mejores resultados que un LIDAR en tareas tales como la detección de peatones. Además de que el alcance de un sistema de visión puede incluso ser superior al de un sensor láser, como se comprobó en el DARPA Grand Challenge (Buelher *et al.* (2005)) con el Vehículo Standley de la universidad de Standford en el que la cámara tenía un alcance de hasta 70 metros contra los 22 metros que tenía el sensor láser.

A continuación se presenta el sistema de visión implementado en un vehículo autónomo de la *Escola de Engenharia de São Carlos* de la *Universidade de São Paulo*. Este algoritmo caracterizará el entorno utilizando flujo óptico. El código tiene como finalidad obtener información del ambiente a partir del movimiento aparente de los objetos en las imágenes. Discriminar objetos en movimiento de aquellos que se encuentran estáticos y determinar el movimiento del vehículo. Este trabajo plantea utilizar la información obtenida para otras funciones tales como la navegación autónoma, exploración del entorno, detección de obstáculos, detección de objetos en movimiento y clasificador de objetos. Debido a que el sistema cuenta con varios sensores con tareas distintas y que demandan muchos recursos computacionales, se requería que el código fuese lo más eficiente posible.

## 2. DESCRIPCIÓN DEL EXPERIMENTO

Los experimentos se realizaron dentro del Proyecto SENA (Sistema Embarcado de Navegación Autónoma), el cual consiste en un sistema enfocado al sensado y a la navegación implementado en un automóvil *Fiat Palio*, que se muestra en la Fig. 1B). El vehículo cuenta con una cámara monocular a color *Infaimon Stingray* mostrada en la Fig. 1A). Este dispositivo captura imágenes con una resolución de 640 x 480 funcionando con una frecuencia de captura de imágenes de 25Hz la cual se encuentra fija al centro del techo en la parte frontal del automóvil.



Figura 1: A) Vehículo donde se realizaron los experimentos B) Cámara monocular Infaimon Stingray

La computadora que realiza el procesamiento es un CPU industrial *Advantech ARK 3440*. Utiliza el sistema operativo Ubuntu 12.04 LTS de 32 bits con una memoria RAM de 7.7 GB, procesador Intel Core i7 2.53GHz x 4. Funciona con

ROS Fuerte (Robot Operating System) y se utilizaron las librerías de visión computacional de OpenCV.

Las pruebas se realizaron en ambientes no controlados en la ciudad de São Carlos y dentro del campus de la universidad bajo distintas condiciones de iluminación y en calles donde la línea del horizonte fuera variable. Además las muestras cuentan con la presencia de otros vehículos, ciclistas y peatones.

El algoritmo describe la dinámica del vehículo urbano autónomo además de la identificación de objetos dinámicos y estáticos en el campo de visión del mismo. Para lograr esto, el algoritmo analiza los puntos característicos de cada una de las imágenes capturadas con la finalidad de rastrearlos y obtener el flujo óptico. Los valores de los vectores de flujo óptico son medidos para luego ser comparados y discriminar las perturbaciones, es decir, vectores que no corresponden a ningún objeto en la imagen. La magnitud y la dirección de los vectores filtrados es medida con la finalidad de determinar cuales corresponden a un objeto en movimiento o si forman parte del ambiente. El flujo óptico causado por los puntos característicos que corresponden al ambiente describirán la dinámica del vehículo, ya que estos vectores son la representación del movimiento aparente del entorno del vehículo. Con la información de dirección, magnitud, la frecuencia de captura de la cámara y la velocidad de procesamiento del algoritmo, se obtiene la velocidad instantánea del vehículo para realizar la odometría. Con esta información también es posible realizar el mapeo del recorrido del mismo. Los vectores que fueron clasificados como objetos dinámicos determinarán una región de interés para que, en proyectos posteriores, serán utilizados para algoritmos clasificadores de elementos tales como vehículos, ciclistas o peatones.

Cuando el vehículo avanza o retrocede, los vectores del campo de flujo óptico aparentemente surgen de un punto en la imagen el cual se conoce como foco de expansión (FoE). Este punto se utiliza para realizar una construcción tridimensional del ambiente además de ayudar a determinar la línea del horizonte, la cual se requiere para conocer la inclinación del camino por el cual transita el vehículo.

A pesar de que se cuentan con sensores de alta precisión como el Velodyne, el IBEO, el IMU y el GPS, implementar un sistema de visión al sistema resulta beneficioso ya que es posible comparar y comprobar los resultados obtenidos con estos sensores con las imágenes obteniendo así una mayor confiabilidad de estos. También es necesario fusionar estos resultados ya que la aplicación requiere un mínimo de errores.

### 3. PROCESAMIENTO DE IMÁGENES

La fase de procesamiento tiene como objetivo mejorar la calidad de la imagen y facilitar la realización de los algoritmos del proyecto. Primeramente se aumentó el contraste para facilitar el reconocimiento de objetos. Después, se le aplicó un filtro no-lineal para reducir el ruido en la imagen y finalmente se transformó a escala de grises. Woods (2008), menciona en su publicación que la respuesta de los filtros no-lineales se basa en ordenar los píxeles contenidos en un área y sustituyendo el valor del píxel central por el obtenido por el filtro. En el presente proyecto se utilizó un filtro bilateral.

De acuerdo con Kaehler (2008), aplicar un filtro reduce aparentemente la resolución de la imagen. Los píxeles tienen la tendencia de ser similares a los píxeles del vecindario. El filtrado, según la definición publicada por Ponce (2003) se refiere a predecir los valores de los píxeles a partir de los valores de sus vecinos. Este se realizó en la imagen colorida ya que el filtro bilateral se efectúa en cada una de las bandas de color de forma separada. Tomasi (1998) demuestra que, utilizando el filtro bilateral, la combinación de las tres bandas ocurre apropiadamente y no aparece el ruido que aparecería si se utilizara un filtro gaussiano convencional. En un suavizado gaussiano ocurriría de manera diferente en cada una de ellas provocando resultados erróneos. Este tipo de filtro, además de reducir el ruido en las imágenes, preserva los bordes como se muestra en la Fig. 2.

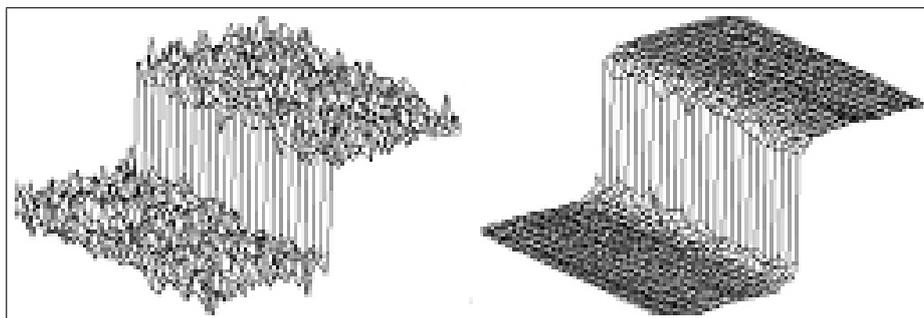


Figura 2: El filtro bilateral elimina el ruido sin afectar los bordes

Esto ocurre ya que las variaciones locales son medidas en cada punto y los valores de los píxeles son promediados con los píxeles vecinos cuyo tamaño y forma dependen de variaciones locales. Una motivación para el suavizado Gaussiano es que los píxeles en una imagen real deben tener pocas variaciones en el espacio y estar relacionados con los píxeles vecinos mientras que el ruido presentará una variación brusca. Este filtro valoriza mejor a los píxeles más similares que a aquellos que presentan más diferencias. El reducir la cantidad de los detalles facilitará el proceso de reconocimiento de puntos característicos. Este filtro se utiliza con una máscara muy pequeña, ya que se requieren los detalles de la imagen para localizar las esquinas y realizar el flujo óptico. El resultado final se muestra en la Fig. 3.



Figura 3: Imagen a la que se le aplicó un filtro bilateral en escala de grises

#### 4. OBTENCIÓN DE PUNTOS CARACTERÍSTICOS

El rastreo de puntos característicos se utiliza en varias aplicaciones de visión computacional de bajo y alto nivel relacionadas con la robótica móvil tales como el Mapeo y Localización Simultanea (SLAM) o la estimación de estructuras a través del movimiento (SfM), como se menciona en el trabajo de Pinz (1994). En el presente proyecto fue necesario encontrar puntos rastreables para conocer la dinámica del automóvil y del entorno del mismo. Si un punto contiene derivadas fuertes en dos direcciones ortogonales, entonces es más probable de que se trate de un punto único, a estos puntos se les llaman esquinas. Hay lugares en la imagen donde la autocorrelación de la matriz de la segunda derivada tiene dos grandes eigenvalores. Lo cual significa que existe una textura o bordes que van en al menos dos direcciones diferentes alrededor de un punto. Las segundas derivadas son útiles ya que no responden a gradientes uniformes. Cuando se consideran solo los eigenvalores de una matriz de autocorrelación, estamos considerando cantidades que son invariables también a la rotación, como lo menciona Kaehler (2008) esto es importante ya que los objetos rastreados también podrían rotar.

Conforme la cámara se mueve, los patrones de intensidad de imagen cambian de manera compleja. Estos cambios son descritos como movimiento de la imagen. Sin embargo Tomasi (1994) especifica que, a pesar del método utilizado para el rastreo, no todas las partes de una imagen contienen información completa sobre el movimiento, a esto se le llama problema de apertura. Para superar esta dificultad, se utiliza el rastreo de esquinas con ventanas de alta frecuencia de contenido espacial o regiones donde una mezcla de derivadas de segundo orden sean suficientemente altas.

Se utilizó el método propuesto por Shi y Tomasi llamado *Good Features to Track* para localizar los puntos característicos. Dicho método surge del trabajo de Harris para la localización de las esquinas. En este método, se considera la matriz de autocorrelación de la segunda derivada de las imágenes en una pequeña ventana alrededor de cada punto. Kaehler (2008) menciona que este método presenta resultados más satisfactorios que el obtener las esquinas con el método de Harris, el cual también es muy utilizado en este tipo de aplicaciones.

En cada una de las imágenes tomadas por la cámara del vehículo se buscan las esquinas, como se muestra en la Fig 4. las cuales, son guardadas en un vector para enseguida ser comparadas con las esquinas de la imagen siguiente. Los vectores que unen cada uno de los puntos en ambas imágenes forman el campo de flujo óptico.

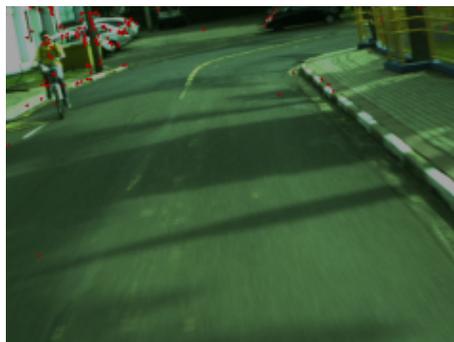


Figura 4: Esquinas de la imagen

#### 5. REPRESENTACIÓN GRÁFICA DEL MOVIMIENTO EN LA IMAGEN

El uso del flujo óptico es uno de los métodos utilizados para la descripción del entorno dinámico de un vehículo. Diferentes técnicas han sido utilizadas para mejorar la generación del flujo óptico en una secuencia de video de los cuales se destacan las comparaciones entre el patrón de flujo, la segmentación de imágenes en movimiento y el tiempo de procesamiento. Existen unos métodos diferenciales como el método propuesto por Lucas y por Kanade (1981), el cual asume como un restricción adicional que el flujo óptico varía ligeramente con respecto a los puntos de objetos vecinos que poseían la misma velocidad y Schunk (1981) asume como una restricción que el flujo óptico está variando con puntos de

objetos vecinos que tiene casi la misma velocidad. Al utilizar cada una de las características de los dos métodos de manera conjunta, las derivadas parciales de Lucas Kanade y el número de iteraciones de Horn Schunck se mejora la segmentación y el flujo óptico obtenido. Estas técnicas por separado o juntas han servido para el desarrollo y mejoramiento de los métodos de obtención del flujo óptico para la identificación de obstáculos con imágenes en movimiento y su respectiva aplicación en la navegación de sistemas robóticos.

El movimiento aparente de los objetos en el plano de una imagen permite comprender la estructura tridimensional del movimiento. Horn (1986) define al flujo óptico como el desplazamiento aparente de los patrones de brillo observados cuando una cámara se mueve con respecto a los objetos capturados. El campo de movimiento es un concepto puramente geométrico. Las bases de la mayoría de los estudios de análisis de movimiento rígido a partir de sistemas de visión se enfoca en el problema del movimiento tridimensional aproximando el campo de movimiento a través del flujo óptico Branca *et al.* (1996). Cuando los objetos se mueven en frente de una cámara o una cámara se mueve a través de un ambiente fijo, existen cambios correspondientes en la imagen y estos pueden ser utilizados para recuperar los movimientos relativos. El campo de movimiento asigna un vector de velocidad a cada punto de la imagen. Idealmente, el flujo óptico corresponde al campo de movimiento sin embargo, el flujo óptico puede ocurrir debido a una fuente de luz en movimiento en un objeto estático. Un punto de partida para la estimación del flujo óptico es asumir que las intensidades de un píxel son transmitidas de una imagen hacia la otra, para realizar esto, se utilizarán las esquinas obtenidas en el punto anterior.

$$I(x, t) = I(x + u, t + 1) \quad (1)$$

En la ecuación (1),  $I(x, t)$  es la intensidad de la imagen en función del espacio  $x = (x, y)^T$  y el tiempo  $t$ .  $u = (u_1, u_2)^T$  es la velocidad bidimensional.

En el presente trabajo se utilizó el método de Lucas-Kanade, el cual trabaja con pirámides de imágenes. Estas son colecciones de imágenes que surgen de una imagen mayor de la cual se extraen muestras de forma consecutiva y de menor tamaño. Collewet (2001) confirma que este algoritmo es adecuado para esta aplicación que discriminación adecuadamente los valores aceptables de aquellos que no son útiles. El algoritmo solo utiliza informaciones locales derivadas de pequeñas ventanas que rodean cada uno de los puntos de interés. La desventaja de utilizar pequeñas ventanas locales es que, cuando se presentan movimientos mayores, pueden mover los puntos fuera de la ventana local y volverse imposibles de encontrar. Este problema es el que lleva al desarrollo del algoritmo piramidal Lucas-Kanade el cual rastrea comenzando desde el nivel más alto de una imagen piramidal con menor detalle y avanzando a imágenes con mayor detalle. Rastrear a través de pirámides de imágenes permite que movimientos largos sean capturados utilizando ventanas.

Una vez obtenidos los vectores de flujo óptico, se medirá cada uno de ellos. A pesar de la buena calidad de los resultados obtenidos, el método de Lucas Kanade no es infalible y puede devolver falsos positivos. Estos vectores son casos aislados que por lo regular tienen una magnitud muy grande, lo cual hace que sea más fácil descartarlos. Se espera que la mayor parte de los puntos característicos detectados provengan de los objetos estáticos, por ejemplo, calles, árboles o vehículos estacionados. La cantidad de vectores que representan los objetos dinámicos tendrán una magnitud diferente a la de los vectores que representan al ambiente. También no surgirán del foco de expansión como el resto de los vectores. Como resultado, se pueden diferenciar las perturbaciones, los objetos dinámicos y los objetos estáticos de la escena. Los resultados de este algoritmo se presentan en la Fig. 5. Con color rojo se representan los objetos dinámicos y con color azul los objetos estáticos. Se considera que vectores individuales aislados no pueden representar objetos tales como peatones o vehículos, por lo que también son descartados y clasificados como falsos positivos. La región por la cual atraviesan los vectores de flujo óptico que han reconocido a un objeto en movimiento es clasificada como una región de interés para utilizarse después como dato de entrada en un clasificador de objetos.

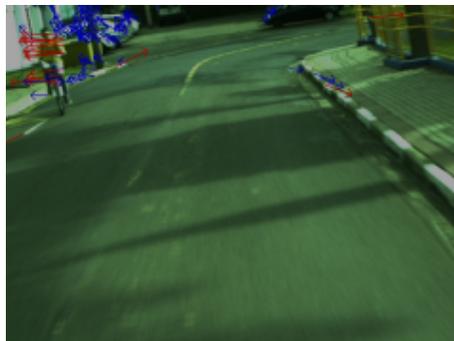


Figura 5: Flujo óptico.

## 6. UBICACIÓN DEL FOCO DE EXPANSIÓN

Según la teoría ecológica de la percepción de Gibson (1986) y posteriormente citada en la investigación de Kuiaski *et al.* (2011), la información óptica del ambiente converge en un solo punto en el espacio formando el arreglo óptico del ambiente dinámico. En este punto, todos los vectores de flujo óptico convergen debido al movimiento de la cámara, por lo que la estimación del FoE proporcionará la información necesaria para describir el movimiento del sensor de visión.

Los vectores más cercanos al FoE correspondientes al ambiente tendrán una magnitud menor mientras que aquellos más alejados de este punto serán mayores. Para poder compararlos, se utilizará el concepto de Vanishing Point mostrado en la Fig. 6. Este punto coincidirá con el FoE, se encuentra al frente del vehículo, teóricamente, en el infinito. Una vez proporcionados los vectores, se mide la magnitud de estos y se obtiene la odometría.

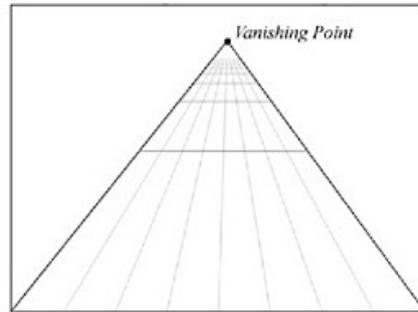


Figura 6: Punto teórico ubicado en el infinito que coincidirá con el foco de expansión del flujo óptico

El vehículo se desplazará la mayor parte del tiempo hacia enfrente, lo cual dará como resultado un campo de movimiento de traslación pura. En este tipo de campo de movimiento, los vectores radian del FoE. En un ambiente en donde no se presente ningún tipo de vibración o elevación, la posición de este punto permanecerá constante en la imagen. Sin embargo, debido a que el vehículo está diseñado para transitar en un ambiente no controlado en donde existieran caminos con frecuentes cambios de elevación, se diseñó un algoritmo que fuera capaz de localizar el FoE ya que en este tipo de situaciones, este cambia constantemente su posición dentro del eje vertical. Para localizar dicho punto, se buscó el lugar del cual aparentemente surgirán los vectores del campo de movimiento y logró ubicarse como se muestra en la Fig. 7. Para lograrlo, se consideró la dirección de los vectores de flujo óptico. El foco de expansión se encontraría en el punto en las columnas y en las filas en donde los vectores de flujo óptico que representan al ambiente cambian de sentido para apuntar a la dirección opuesta en un ángulo cercano a los 180°. Este principio también puede ser utilizado para determinar magnitud de los cambios de dirección del vehículo.



Figura 7: Ubicación de la línea del foco de expansión

## 7. ESTIMACIÓN DE LA DIRECCIÓN DEL VEHÍCULO

Para obtener los parámetros tridimensionales se analiza el campo de vectores en dos dimensiones. Asumiendo un modelo de proyección perspectiva en el cual un punto  $P = (X, Y, Z)$  se proyecta en la imagen en el punto  $(x, y) = f(\frac{X}{Z}, \frac{Y}{Z})$  en donde  $f$  es la distancia focal. En el trabajo de Longuet-Higgins y Prazdny Prazdny (1980) se derivan las siguientes ecuaciones para describir el movimiento rígido general de un observador dentro de un mundo estacionario.

$$u = \frac{T_x - xT_z}{Z(x, y)} - xyR_x + (1 + x^2R_y - yR_z) \quad (2)$$

$$v = \frac{T_y - yT_z}{Z(x, y)} - (1 + y^2)R_x + xyR_y + yR_z \quad (3)$$

Donde  $(T_x, T_y, T_z)$  son las componentes tridimensionales de la velocidad translacional,  $(R_x, R_y, R_z)$  las componentes tridimensionales rotacionales de la velocidad,  $(u, v)$  las velocidades proyectadas de un punto  $(X, Y, Z)$  en una imagen. Se consideró que la mayor cantidad de vectores obtenidos provendrá del ambiente, por lo que los valores superiores corresponderán a objetos dinámicos. Aquellos que presentan velocidades muy superiores o muy inferiores a la media, serán descartados e identificados como perturbaciones. Los resultados obtenidos se muestran en la Fig. 8.



Figura 8: (Derecha) Objetos estáticos. (Izquierda) Objetos dinámicos.

## 8. RESULTADOS

Las imágenes obtenidas son procesadas para obtener una mejor calidad en los puntos característicos. Filtrando las perturbaciones presentes en las imágenes se disminuyó la variación de la posición de los puntos característicos y por consecuencia se aumentó la calidad del flujo óptico. Si las imágenes son afectadas por un ruido uniforme, el código no será capaz de identificar correctamente un punto característico en dos imágenes diferentes. El algoritmo localizó a los objetos dinámicos debido a que la magnitud de los vectores de flujo óptico producidos por dichos objetos era significativamente diferente de aquellos que eran producidos por movimiento aparente del ambiente. Estos vectores discriminados proporcionan la información suficiente para indicar una región de interés y, en proyectos posteriores, analizarla en un clasificador de objetos.

En las Figuras 9, 10, 11 y 12, los vectores correspondientes a objetos dinámicos aparecen de color rojo mientras que aquellos correspondientes al ambiente aparecen de color azul. Los objetos dinámicos son descartados del campo de flujo óptico y, a partir de las magnitudes de los vectores del campo de movimiento restantes, es decir aquellos que describen el movimiento aparente del ambiente, se determina la magnitud del movimiento del vehículo, su dirección, la odometría y su velocidad.

Entre los objetos dinámicos que se encuentran dentro del entorno del vehículo y al alcance de la cámara están otros vehículos, motociclistas y ciclistas. El algoritmo por si solo no es capaz de determinar que clase de objeto es el que está detectando, sin embargo detecta la presencia de objetos que independientemente de su naturaleza requieren ser esquivados. En las pruebas realizadas se logró identificar la presencia de objetos en movimiento tales como automóviles tal como se muestra en las Figuras 9, 10, 11 y motociclistas, mostrados en la Figura 12



Figura 9: Vehículos en movimiento reconocidos por el algoritmo. Imágenes tomadas dentro del campus de la EESC-USP.

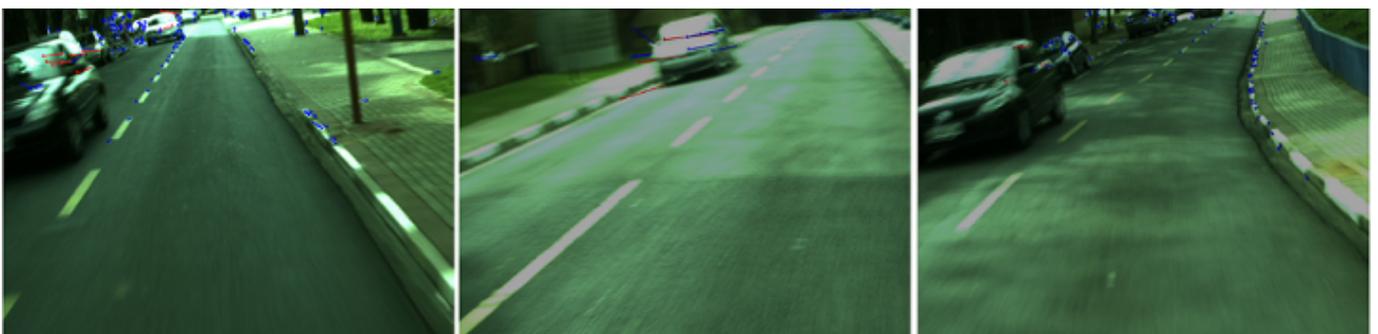


Figura 10: Automóviles. Objetos dinámicos reconocidos por el algoritmo. Imágenes tomadas dentro del campus de la EESC-USP.

Algunas de las imágenes capturadas resultan demasiado borrosas debido a la velocidad del vehículo y la velocidad de captura de la cámara. La baja calidad de dichas imágenes complicaría su identificación si se utilizara un algoritmo

especializado para reconocer la morfología del automóvil, sin embargo, esa no fue una limitación para que el sistema del presente trabajo fuera capaz de detectar la presencia de un objeto dinámico en la escena como se muestra en las imágenes de la Figura 11.



Figura 11: Automóviles. Objetos dinámicos reconocidos por el algoritmo a pesar de que la imagen aparece borrosa debido a la velocidad de los objetos presentes.

Dentro de las pruebas realizadas también se contó con la presencia de motociclistas cuya presencia fue detectada por el algoritmo.



Figura 12: Motocicletas. Imágenes tomadas en la ciudad de Sao Carlos bajo distintas condiciones de iluminación.

El horizonte se ubicó utilizando el foco de expansión de los vectores de flujo óptico. La posición de dicho punto es variable debido a los cambios de elevación de las calles por las cuales transita el vehículo. En la figura 13 se muestra posición distintas en las cuales se encontró el horizonte. Este punto ayuda a tener más información sobre el ambiente en el cual está transitando el vehículo. Se requiere para interpretar correctamente los vectores de flujo óptico del campo de movimiento.



Figura 13: Ubicación de la línea del foco de expansión obtenido por medio de la dirección de los vectores de flujo óptico.

## 9. CONCLUSIONES

El ser humano no está exento a cometer equivocaciones al conducir. Debido a esto es importante diseñar aplicaciones que puedan disminuir las altas cifras de accidentes viales. Los vehículos autónomos se encontrarán presentes en un futuro no muy lejano y requieren de sistemas con una alta confiabilidad que garanticen la seguridad del usuario. Entre los métodos utilizados para realizar tareas tales como navegación, detección de obstáculos, medición de velocidad, dirección y la odometría, los sistemas de visión representan una alternativa de la cual se obtienen resultados muy aceptables a un costo relativamente bajo.

El algoritmo no tiene como finalidad el identificar un objeto dinámico de otro sino de identificar la presencia de estos en el entorno del vehículo autónomo. Los vectores de flujo óptico que representan a los objetos dinámicos no neces-

riamente delimitan la región en la cual se encuentra dicho objeto. Es posible que muchos de los puntos característicos que delimitarían al objeto se pierdan debido a la velocidad y resolución de la cámara así como por las condiciones de iluminación. Se debe considerar ampliar la región de interés al rededor de los vectores para luego utilizar esa región para la correcta clasificación del objeto dinámico detectado.

A pesar de que los resultados parciales demuestran la funcionalidad del proyecto, aun es posible mejorar el trabajo utilizando una cámara de mayor resolución, adecuando el ángulo de inclinación de la cámara para obtener un mayor campo de visión. El siguiente problema a resolver es el determinar una región de interés a partir de la región en donde se encuentran los objetos dinámicos para ingresar estos valores a un algoritmo clasificador de objetos. Los resultados serán fusionados con los obtenidos por otros sensores lo cual aumentará la confiabilidad de los datos obtenidos.

## 10. AGRADECIMENTOS

Al Laboratorio de Robotica de la Universidad Autónoma de Ciudad Juárez, Chihuahua, México, en especial al Dr. Edgar Martinez García. Al laboratorio de dinámica de la Escola de Engenharia de São Carlos y a los miembros del equipo del Proyecto SENA.

## 11. REFERENCIAS

- Branca, A., Stelle, E. and Distant, A., 1996. "Passive navigation using focus of expansion". *Applications of Computer Vision, 1996. WACV '96., Proceedings 3rd IEEE Workshop on*, pp. 64–69.
- Broggi, A., Grisleri, Z. and Paolo, Z., 2013. "Sensors technologies for intelligent vehicles perception systems: a comparison between vision and 3d-lidar". *16th International IEEE Annual Conference on Intelligent Transportation System (ITSC 2013), The Hague, The Netherlands*, pp. 887–892.
- Buelher, M., Sanjiv, S. and Iagnemma, K., 2005. *The 2005 DARPA Grand Challenge*. Springer Tracts in Advanced Robotics., Berlin Heidelberg. 428 p.
- Buelher, M., Sanjiv, S. and Iagnemma, K., 2009. *The DARPA Urban Challenge*. Springer Tracts in Advanced Robotics., Berlin Heidelberg. 428 p.
- Collewet, C.K.C., 2001. "Improving feature tracking by robust points of interest selection". *Vision Modeling and Visualization Conference*, pp. 415–422.
- DARPA, 2014. "Darpa grand challenge". [http : //www.darpa-grandchallenge.com/](http://www.darpa-grandchallenge.com/).
- Form, J.M.W.T., 2008. "Realizing complex autonomous driving maneuvers the approach taken by team carolo at the darpa urban challenge." *Vehicular Electronics and Safety*, pp. 232–236.
- Gibson, J.J., 1986. *The Ecological Approach to Visual Perception*. Lawrence Erlbaum Associates, Hillsdale.
- Horn, B.K.P., 1986. *Robot Vision*. The MIT Electrical Engineering and Computer Science Series., Boston, Massachusetts. 428 p.
- Kaehler, G.B.A., 2008. *Learning OpenCV, Computer Vision with the OpenCV Library*. O'Reilly, Sebastopol, California. 428 p.
- Kanade, B.D.L.T., 1981. "An iterative image registration technique with an application to stereo vision". *International Joint Conferences on Artificial Intelligence*, , No. 81, pp. 674–679.
- Kuiaski, J.R., Lazzaretti, A.E. and Neto, H.V., 2011. "Focus of expansion estimation for motion segmentation from a single camera". *VII Workshop de Visão Computacional – WVC 2011*, pp. 272–277.
- Niester, D., Naroditsky, O. and Bergen, J., 2004. "Visual odometry". *IEEE International Computer Society Conference on Computer Vision and Pattern Recognition*, , No. 3, pp. 332–340.
- Pinz, C.F.A., 1994. "Spatio-temporal good features to track". *ICCV2013 Workshop paper*, pp. 246–253.
- Ponce, D.F.J., 2003. *Computer Vision*. Pearson Prentice Hall, Upper Saddle River, New Jersey. 428 p.
- Prazdny, H.C.L.H.K., 1980. "The interpretation of a moving retinal image". *The Royal Society, London*, pp. 232–236.
- Schuck, B.K.H.B.G., 1981. "Determining optical flow". *Technical Symposium East. International Society for Optics and Photonics.*, pp. 319–331.
- Spectrum, I., 2010. "Googles autonomous car takes to the streets". [http : //spectrum.ieee.org/automaton/robotics/artificial-intelligence/googles-autonomous-car-takes-to-the-streets](http://spectrum.ieee.org/automaton/robotics/artificial-intelligence/googles-autonomous-car-takes-to-the-streets).
- Tomasi, J.S.C., 1994. "Good features to track". *IEEE Conference on Computer Vision and Patter Recognition*, , No. 3, pp. 332–340.
- Tomasi, R.M.C., 1998. "Bilateral filtering for gray and color images". *IEEE International Conference on Computer Vision, Bombay, India*, , No. 3, pp. 332–340.
- WHO, 2013. "Road traffic injuries". [http : //www.who.int/mediacentre/factsheets/fs358/en/](http://www.who.int/mediacentre/factsheets/fs358/en/).
- Woods, R.G.R., 2008. *Digital Image Processing*. Pearson Prentice Hall, Upper Saddle River, New Jersey. 428 p.

## Use of optical flow to describe the dynamic environment of an urban vehicle

Genaro Ellian Herrera Bandin, [ellianhbandin@gmail.com](mailto:ellianhbandin@gmail.com)<sup>1,2</sup>

João Pedro Finoto, [jpfm2210@gmail.com](mailto:jpfm2210@gmail.com)<sup>1</sup>

Ingrid Lorena Argote Pedraza, [ingridargote@usp.br](mailto:ingridargote@usp.br)<sup>1</sup>

Prof. Marcelo Becker, [becker@sc.usp.br](mailto:becker@sc.usp.br)<sup>1</sup>

<sup>1</sup>Escola de Engenharia de São Carlos - Universidade de São Paulo, Av. Trabalhador São-carlense, 400, Pq Arnold Schimidt São Carlos - SP/Brasil, CEP 13566-590

<sup>2</sup>Universidad Autónoma de Ciudad Juárez, Av. Plutarco Elías Calles #1210 , Fovissste Chamizal , Ciudad Juárez, Chihuahua, México , C.P. 32310

**Abstract:** *This paper is aimed to present an algorithm that describes the dynamic environment of an urban autonomous vehicle which interprets the data from the optical flow of the scene. The tests were done as part of the SENA Project (Autonomous Embedded Navigation System) of the São Carlos Engineering School - University of São Paulo. The system has a color monocular camera working with images of 640x480 pixels and with a frequency of 25fps. The algorithm identifies the corners by the Shi-Tomasi method. Data of the first image is stocked to compare it with the data of the next image and afterwards the optical flow is obtained with the Lucas-Kanade method. The orientation and the length of the vectors of the optical flow are discriminated in order to determinate whether the object is dynamic or static. Given the acquired information coming from the vectors of the static objects (environment), the algorithm obtains the orientation and speed of the vehicle. The algorithm determines the focus of expansion from the optical flow and with this information, the horizon of the image can be located, and the car's tridimensional speed can be calculated.*

**Keywords:** *Autonomous Vehicles, Computer Vision, Optical Flow, Visual Odometry*