



## INFORME CIENTIFICO DE BECA

Legajo N°:

**TIPO DE BECA** Beca de Estudio

**PERIODO** 2016-2017

### 1. DATOS PERSONALES

*APELLIDO: Bussi*

*NOMBRES: Ulises*

*Dirección Particular: Calle:*

*Localidad: Ezpeleta CP: 1882 Tel:*

*Dirección electrónica (donde desea recibir información, que no sea "Hotmail"):  
ulises.bussi@unq.edu.ar*

### 2. TEMA DE INVESTIGACION (Debe adjuntarse copia del plan de actividades presentado con la solicitud de Beca)

Visión Artificial para seguridad ciudadana

**PALABRAS CLAVE (HASTA 3)** Visión Seguridad Camaras

### 3. OTROS DATOS (Completar lo que corresponda)

**BECA DOCTORAL 1° AÑO** (ex ESTUDIO 1° AÑO): *Fecha inicio:* 1 de Abril de 2016

**BECA DOCTORAL 2° AÑO** (ex ESTUDIO 2° AÑO): *Fecha inicio:*

**BECA DOCTORAL 3° AÑO** (ex PERFECCIONAMIENTO 1° AÑO): *Fecha inicio:*

**BECA DOCTORAL 4° AÑO** (ex PERFECCIONAMIENTO 2° AÑO): *Fecha inicio:*

### 4. INSTITUCION DONDE DESARROLLA LA TAREA

*Universidad y/o Centro: Universidad Nacional de Quilmes*

*Facultad:*

*Departamento: Departamento de Ciencia y Tecnología*

*Cátedra:*

*Otros:*

*Dirección: Calle: Roque Saenz Peña N°: 352*

*Localidad: Bernal CP: 1876 Tel: 011-4356-7100*

### 5. CARGO UNIVERSITARIO (si existe, especificar categoría, dedicación, condición de ordinario, regular o interino):

Profesor instructor interino con dedicación simple desde Septiembre del 2016

### 6. CARGOS EN OTRAS INSTITUCIONES:

### 7. DIRECTOR DE BECA

*Apellido y Nombres: Safar, Felix Gustavo Emilio*

*Dirección Particular: Calle:*

Localidad: M.B. Gonnet CP: 1897 Tel:

Dirección electrónica: HYPERLINK "mailto:felix.safar@unq.edu.ar" [felix.safar@unq.edu.ar](mailto:felix.safar@unq.edu.ar)

HYPERLINK ""

[felixsafar@gmail.com](mailto:felixsafar@gmail.com)

## 8. RESUMEN DE LA LABOR QUE DESARROLLA

*Descripción para el repositorio institucional. Máximo 150 palabras.*

El objetivo general de este proyecto es la generación de conocimiento aplicado para el desarrollo de sistemas autónomos basados en técnicas de Visión e Inteligencia Artificial para el procesamiento de escenas visuales dinámicas. Estos sistemas asistirán a los operadores humanos en tareas de análisis, toma de decisión y control motor.

Los objetivos específicos de este proyecto apuntan al desarrollo de un sistema autónomo basado en Visión Artificial para tareas de Seguridad Ciudadana. Estos objetivos consisten en una serie de sub-problemas concatenados: **(A)**- Desarrollo de un sistema de adquisición de imágenes compuesto por una cámara con visión de campo amplio (*fish-eye*) y una cámara *Pan-Tilt-Zoom* (PTZ). **(B)** Implementación de algoritmos para la clasificación de videos en *background* (BKG) y *foreground* (FRG) para la detección de objetos en movimiento. **(C)** Algoritmos para el seguimiento de objetos móviles. **(D)** Algoritmos de clasificación de los objetos y sus patrones de actividad. **(E)** Algoritmo de posicionamiento y nivel de zoom de la cámara PTZ para adquirir información de un objeto de interés en una resolución mayor a la obtenida por la cámara *fish-eye*.

## 9. EXPOSICION SINTETICA DE LA LABOR DESARROLLADA EN EL PERIODO.

*Debe exponerse la orientación impuesta a los trabajos, técnicas empleadas, métodos, etc., y dificultades encontradas en el desarrollo de los mismos, en el plano científico y material. Si corresponde, explicita la importancia de sus trabajos con relación a los intereses de la Provincia.*

**Tareas desarrolladas en relación al objetivo (A): Estudio de teoría, modelado y simulación del sistema de cámaras**

En el año correspondiente a la beca, y en el marco del grupo de Visión Artificial y Robótica del Programa de Investigación "Estrategias de ingeniería en automatización, computación y procesos industriales aplicadas a la resolución de problemas tecnológicos" de la Universidad Nacional de Quilmes, cuyo Director es Félix Safar, se comenzó con el estudio de la teoría asociada a cámaras VCA y PTZ [1,2,3]. En la actualidad, tanto organizaciones privadas como en el ámbito público, es común el uso de sistemas de video-vigilancia para garantizar la seguridad de las personas o bienes de valor. Estos sistemas trabajan actualmente en su mayoría con cámaras de proyección planas de tecnología IP (*internet protocol*). Las cámaras de proyección plana pueden ser fijas o del tipo PTZ, ambas presentan una limitación fundamental, tienen un campo de visión máximo acotado a 60°x60° lo que no permite tener un panorama general y es necesario re-orientar la cámara hacia los sitios de interés. Las cámaras PTZ permiten realizar esta orientación constantemente gracias a un sistema de motores y la capacidad de manejo remoto de las mismas. En ambos casos la re-orientación constante de las cámaras genera un desgaste debido a piezas móviles.

En los trabajos estudiados se plantea la posibilidad de utilizar cámaras de campo amplio, ya que permiten obtener una información panorámica global, con un campo aproximado de 360°x180. Estas cámaras sin partes móviles permiten observar en varias direcciones al mismo tiempo. Como desventaja este tipo de cámaras presenta deformaciones de la imagen donde una trayectoria recta en el espacio se transforma en una curva. Otra desventaja particular es que al tratarse de una cámara que toma imágenes en un campo tan amplio, al enfocarse en una región de interés reducida la definición disponible disminuye drásticamente.

Se implementó, estudió y simuló, un modelo de cámara PTZ y uno de VCA que incorpora los factores intrínsecos (relacionados con óptica del sistema); como también los extrínsecos (relacionados con la posición y orientación de la cámara) . Para la cámara PTZ se utilizó el modelo clásico de *pin hole* de cámaras planas permite establecer una relación entre la posición tridimensional de un objeto en el espacio y su proyección sobre la imagen obtenida por la cámara.

Para el estudio y simulación de la cámara VCA se utilizó el modelo estereográfico, este modelo se basa en plantear los puntos del espacio 3D en coordenadas esféricas, que se proyectan al plano de la imagen. Dicha proyección depende de una constante multiplicativa "k" que define un factor de escalamiento de la imagen [4].

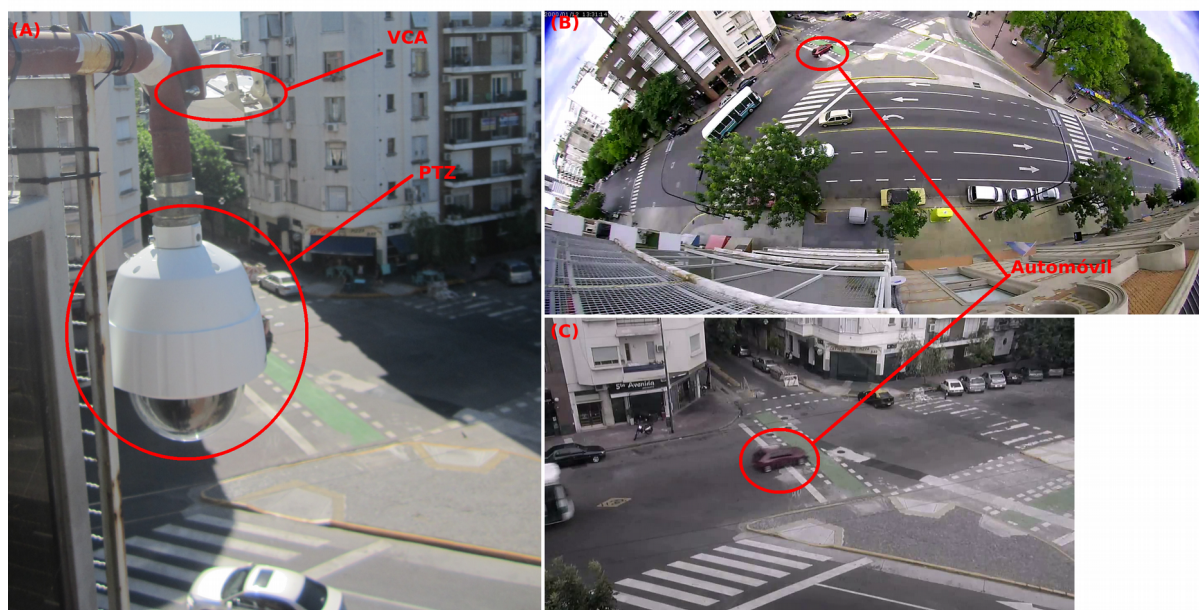
La simulación del sistema de adquisición de imágenes permitió comprender la geometría del sistema y a partir de los modelos se comenzó a trabajar en la identificación de los parámetros.

### Actividades relacionadas con la implementación del sistemas

En base a lo expuesto anteriormente se propuso utilizar en conjunto las dos cámaras para obtener información general con la cámara de campo amplio y encontrar eventos de interés y su dirección, para luego enfocar una PTZ en la dirección establecida anteriormente.

Es común que los sistemas de video vigilancia estén conformados por cámaras y software de distintos tipos y marcas, ya sea por utilidad, precio, o disponibilidad. tanto a la hora de diseño de un sistema desde cero o en la modificación para ampliación del mismo. Este hecho trae como inconveniente la necesidad de un software para administración y configuración especializado para cada marca y modelo.

Como solución a dicho inconveniente se creó ONVIF (*Open Network Video Interface Forum*), un estándar abierto utilizado como interfaz de productos de seguridad sobre IP, basado en *web services*. Este sistema permite la configuración, detección, administración y control de varios dispositivos de distintas marcas. Se estudió el protocolo y se realizó un programa en *python* para acceder, controlar y adquirir con las cámaras PTZ y VCA en colaboración con *Lilian García* y *Sebastián Arroyo*, ambos también integrantes del Programa de Investigación UNQ antes mencionado. Se diseñaron y llevaron a cabo experimentos para la calibración de los parámetros intrínsecos de los modelos descritos en la sección anterior, como ser: la distancia focal de la cámara PTZ y el parámetro "k" de la cámara VCA. Se diseñaron y llevaron a cabo experimentos para la adquisición conjunta utilizando los programas previamente mencionados donde se aplica el protocolo estudiado. Se obtuvieron así videos de tráfico de autos para su posterior análisis (ver figura 1).



**Figura 1:** (A) esquema de montaje de las cámaras VCA y PTZ. (B) imagen global adquirida con la cámara de campo amplio. (C) imagen enfocada de la cámara PTZ en mismo tiempo.

### **Actividades relacionadas con los objetivos (B,C) del proyecto: Detección y seguimiento de objetos en movimiento.**

Se trabajó en la implementación del algoritmo *Background-Foreground* para separar las regiones con movimiento de los videos de las estáticas. Con el fin de detectar los objetos móviles.

Se estudió la teoría de filtro de Kalman para realizar estimaciones tanto de la posición como la velocidad de los objetos. También se estudió la posibilidad de utilizar fuentes distintas de información para la estimación (fusión de sensores) y la extensión no lineal del filtro (EKF). Se probaron 4 modelos para estimar sobre una serie de puntos obtenidos de una trayectoria de un automóvil en una de las filmaciones: 1-partícula libre con velocidad constante, 2-modelo de desplazamiento de la bicicleta con velocidad constante, 3-partícula libre con aceleración constante, 4-modelo de la bicicleta con aceleración constante. Estos 4 modelos tienen 4, 5, 6 y 7 estados cada uno. Quedó pendiente la realización de un experimento donde se pueda obtener información precisa de la velocidad de un automóvil para validar las estimaciones de velocidad de estos modelos y asignarles un grado de precisión y exactitud a cada uno.

### **Actividades relacionadas con el objetivo (D) del proyecto: Clasificación de los objetos**

Para la detección y clasificación, se decidió por estudiar redes neuronales convolutivas, ya que esta herramienta presenta una gran variedad de usos y muy buenos resultados. Este tipo de algoritmos forma parte del estado del arte [5].

Las redes neuronales son un tipo de algoritmo computacional cuyo funcionamiento tiene una base biológica. Las redes convolutivas cuentan con varias capas donde en cada una de ellas se realiza una convolución entre la entrada y una matriz de pesos. Estos algoritmos presentan un muy buen desempeño a la hora de clasificar (poner una etiqueta a una imagen), detectar (ubicar un objeto) y segmentar (separar distintas clases a lo largo de la imagen). Como contrapartida, este tipo de redes suele tener una gran cantidad de capas con varios núcleos cada una, por lo cual, la cantidad de operaciones y memoria necesaria por imagen es alta y también lo es la cantidad de parámetros que lleva la red ( del orden de millones y miles de millones).[6]

A pesar de esto, el exponencial crecimiento de las capacidades de las computadoras en los últimos años, y la capacidad de realizar operaciones en los GPUs (unidades de procesamiento gráfico) ha permitido la implementación de estas redes, que utilizan principalmente estos últimos, por su velocidad para realizar operaciones como la convolución.

Se estudió en detalle la bibliografía y el funcionamiento de estas redes, y se probaron algunas herramientas de programación para placas de la marca NVIDIA, utilizando las librerías de CUDA. Se realizaron pruebas con librerías de python como Scikitlearn (*Machine Learning*) [7], Theano (*Large scale computation and GPU use*)[8], Keras (Deep Learning) [9] y se pudieron implementar algunos casos sencillos de clasificación y pruebas de detección en pequeña escala con el hardware disponible de la universidad (Geforce GTX 980 con 2048 núcleos), sin embargo, limitaciones y problemas de compatibilidad de la plataforma GPU impedían aumentar la complejidad y tamaño de las redes para las tareas deseadas, retrasaron significativamente el avance del trabajo en este tema específico durante meses.

Una vez solucionado el problema, se analizó el uso de algoritmos más complejos para realizar la detección. Se realizaron pruebas sobre bases de datos de prueba disponibles en Internet como las de números manuscritos de la base de datos MNIST y se logró plena comprensión del funcionamiento de estas redes. Quedan para desarrollar a futuro las pruebas sobre datos obtenidos en los experimentos de grabación descritos anteriormente.

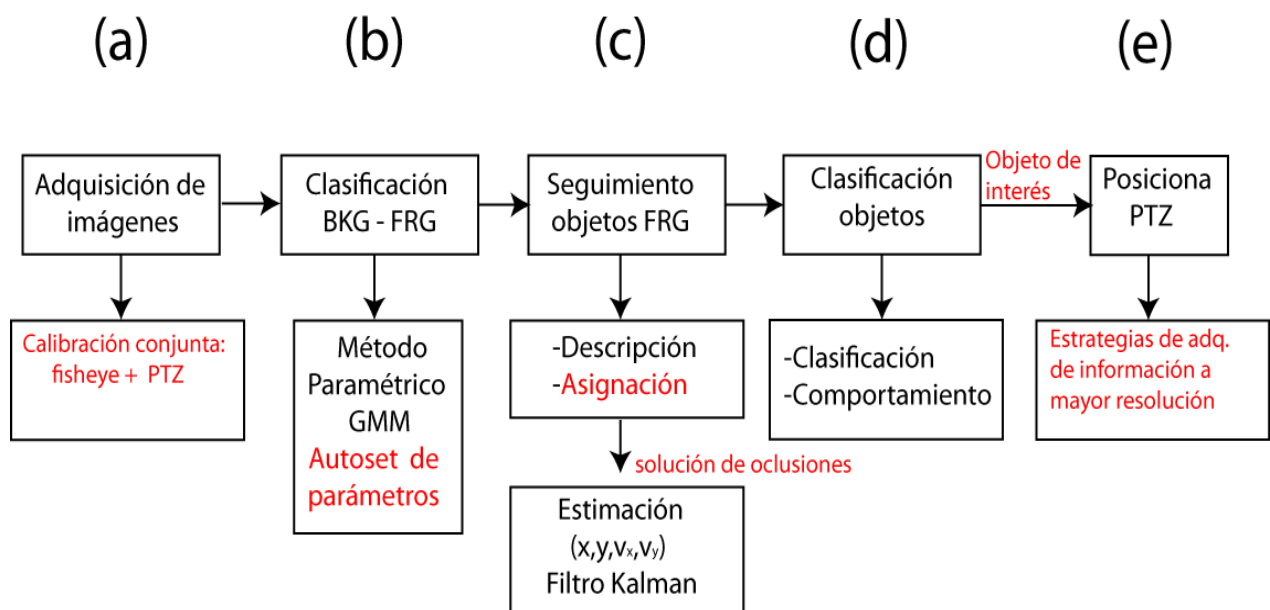
- [1] Arroyo, S. I., Safar, F., & Oliva, D. (2015, October). Georeferenced feature tracking in wide field images. In *Information Processing and Control (RPIC), 2015 XVI Workshop on* (pp. 1-6). IEEE.
- [2] Kannala, J., & Brandt, S. S. (2006). A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. *IEEE transactions on pattern analysis and machine intelligence*, 28(8), 1335-1340.
- [3] Arroyo, S. I., Safar, F., & Oliva, D. (2016, June). Probabilidad de infracción de velocidad de vehículos utilizando visión artificial en cámaras de campo amplio. In *Biennial Congress of Argentina (ARGENCON), 2016 IEEE* (pp. 1-6). IEEE.
- [4] Stanganelli, D., Oliva, D. E., Noblia, M., & Safar, F. (2014, June). Calibración de una cámara fisheye comercial con el modelo unificado para la observación de objetos múltiples. In *Biennial Congress of Argentina (ARGENCON), 2014 IEEE* (pp. 147-152). IEEE.
- [5] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105).
- [6] Deep Learning, Ian Goodfellow and Yoshua Bengio and Aaron Courville, MIT Press, deeplearningbook.org, 2016.
- [7] Scikit-learn.org, scikit learn library
- [8] Deeplearning.net , theano library
- [9] Keras.io , keras library

**Plan de Trabajo: Desarrollo de un sistema autónomo basado en Visión Artificial con aplicación al ámbito de la Seguridad Ciudadana.**

**Postulante:** Ulises Bussi  
**Director:** MSc. Ing. Félix Safar  
**Codirector:** Dr. Damián Oliva

**1-Objetivo general del Plan de Trabajo:** El objetivo general de este plan es la generación de conocimiento aplicado para el desarrollo de sistemas autónomos basados en técnicas de Visión e Inteligencia Artificial para el procesamiento de escenas visuales dinámicas. Estos sistemas asistirán a los operadores humanos en tareas de análisis, toma de decisión y control motor.

**2-Objetivos específicos:** Los objetivos específicos de este plan apuntan al desarrollo de un sistema autónomo basado en Visión Artificial para tareas de Seguridad Ciudadana. Los objetivos específicos, se presentan como una serie de sub-problemas concatenados (ver Figura 1): **(a)**- Desarrollo de un sistema de adquisición de imágenes compuesto por una cámara con visión de campo amplio (*fisheye*) y una cámara *Pan-Tilt-Zoom* (PTZ). **(b)** Implementación de algoritmos para la clasificación de videos en *background* (BKG) y *foreground* (FRG) para la detección de objetos en movimiento. **(c)** Algoritmos para el seguimiento de objetos móviles. **(d)** Algoritmos de clasificación de los objetos y sus patrones de actividad. **(e)** Algoritmo de posicionamiento y nivel de zoom de la cámara PTZ para adquirir información de un objeto de interés en una resolución mayor a la obtenida por la cámara *fisheye*.



**Figura 1: Sistema propuesto:** Cada objetivo del plan está asociado a uno de los módulos (a-e). En rojo se resaltan los problemas de mayor complejidad. El plan descrito contempla el desarrollo completo en el marco de una tesis doctoral.

**3- Introducción y Antecedentes:** El desarrollo tecnológico actual produce un flujo de información digital que crece exponencialmente en el tiempo y es difícil de interpretar por un número (siempre limitado) de operadores humanos. Por este motivo, es fundamental el desarrollo de sistemas autónomos que asistan a los operadores

humanos en el análisis, toma de decisión y eventualmente, la generación de acciones motoras en función de la información adquirida. En los últimos años se produjo una reducción drástica en el costo del hardware para adquisición y procesamiento de imágenes. Al mismo tiempo surgieron nuevos algoritmos computacionales que revolucionaron el análisis dinámico de la información visual que se encuentran disponibles a través de librerías abiertas como *OpenCV*. Estos hechos convergen en la posibilidad de procesar grandes volúmenes de información visual en tiempo real para implementar soluciones de procesamiento automatizadas con costos moderados [1].

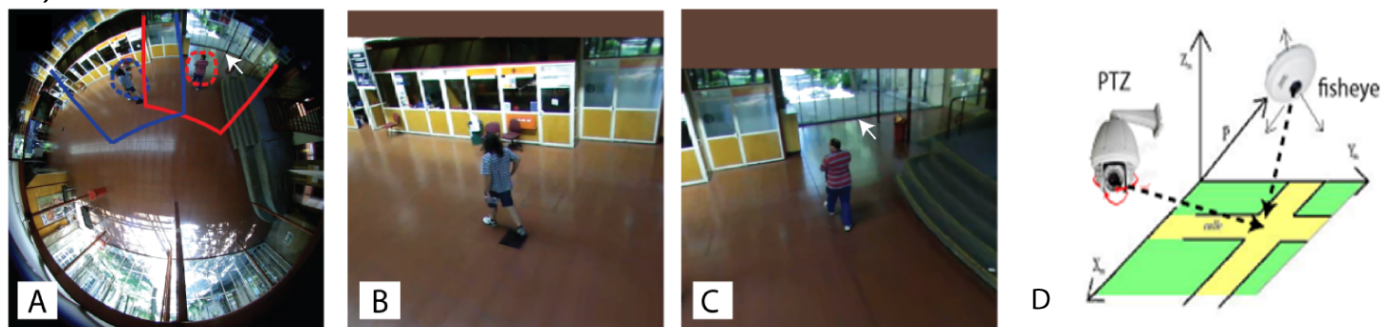
Actualmente las tareas de *video detección* (VD) de escenas dinámicas tienen relativamente bajo nivel de automatización pese a las notables capacidades de adquisición de datos y de cómputo, existe un amplio campo de investigación y desarrollo para estas aplicaciones. Estas consideraciones conducen al objetivo principal de este plan que consiste en la investigación y el desarrollo a nivel prototipo experimental de un sistema de *video-detección autónoma* (VDA) para imágenes de *visión de campo amplio* (VCA) con el propósito de asistir a la VD realizada por operadores humanos en tareas de análisis, toma de decisión y control motor.

El plan está orientado a mejorar dos aspectos fundamentales de la VD. En el objetivo (a) se propone utilizar técnicas que aumenten el *campo de visión* (CV) manteniendo la posibilidad de analizar regiones de interés en alta resolución. Por su parte, los objetivos (b) a (d) se orientan a la aplicación y mejora de algoritmos autónomos que asistan a los operadores humanos en el análisis de los videos. El objetivo (e) se orienta al comando motorizado de una cámara del tipo PTZ (*Pan-Tilt-Zoom*) para seguimiento automático en alta resolución de los objetos detectados por el sistema VCA.

La estrategia que se propone seguir posee un amplio espectro de aplicaciones potenciales y en este proyecto nos concentraremos en el ámbito estratégico de la Seguridad Ciudadana [5]. A continuación se describe el estado del arte en la investigación de cada uno de los módulos que requiere el sistema VDA+VCA.

**3.a - Sensores. Sistema de Visión de Campo Amplio:** Una de las innovaciones para el sistema de VDA será la utilización de sensores que permitan aumentar el campo de visión a todo un hemisferio ( $360 \times 180^\circ$ ), utilizando cámaras con lentes del tipo ojo de pez (cámaras *fisheye* [6]). En los sistemas actuales de video detección (VD), para observar una escena amplia con resolución razonable se utilizan varias cámaras fijas y/o una cámara PTZ operada por un humano con el objetivo de observar las regiones de interés en mayor resolución. Las cámaras PTZ tienen un campo visual máximo de aproximadamente  $60^\circ \times 60^\circ$  por lo que un usuario debe reorientar mecánicamente la cámara para adquirir las regiones de interés perdiendo momentáneamente la visión en otras regiones. Dada esta limitación, es natural preguntarse si pueden usarse sistemas con visión de campo amplio (VCA), como cámaras *fisheye*, que permiten medir un hemisferio de la escena aproximadamente  $360^\circ \times 180^\circ$  [6]. Con este aumento de campo visual, pueden abarcarse grandes extensiones de espacio. Para mantener una relación razonable de píxeles por metro cuadrado estas cámaras utilizan sensores de alta resolución (alrededor de 5 mega píxeles) (ver figura 2A). Estas cámaras no tienen partes móviles y permiten observar simultáneamente en muchas direcciones. En el trabajo [2] presentamos nuestros primeros desarrollos de un sistema de visión de campo amplio para el seguimiento de múltiples objetos en la escena visual. Para esto calibramos los parámetros intrínsecos de una cámara *fisheye* comercial y mostramos que el proceso de formación de imágenes puede ajustarse con el Modelo Unificado. Utilizamos este modelo para corregir las distorsiones provocadas por la cámara *fisheye* y simulamos imágenes adquiridas por una cámara de proyección plana en cualquier dirección del campo visual. Finalmente propusimos un algoritmo simple para detectar regiones móviles en la imagen panorámica, seguir estas regiones de interés a través del tiempo y corregir las distorsiones a medida que el objeto se mueve en la escena visual.

En la figura 2 se puede observar que, aunque la cámara *fisheye* puede observar un panorama amplio (y así detectar todos los objetos de interés en la escena); su aplicación para tareas de video vigilancia, tiene una limitante en la resolución de las imágenes adquiridas. Por este motivo, en este proyecto nos proponemos desarrollar un sistema de mayor complejidad que coordine la información VCA de la cámara *fisheye* y dirija una cámara PTZ para obtener información en mayor resolución de los objetos de interés. La solución de esta problemática determina el primer objetivo de este plan donde nos proponemos una calibración conjunta de los dos tipos de cámaras (figura 2D).



**Figura 1: Sistema de Video detección con una cámara *fisheye*.** En la figura 2A se ve un ejemplo de la imagen obtenida por la cámara *fisheye*. La cámara observa un hemisferio completo pero introduce distorsiones importantes. El algoritmo que implementamos detecta las regiones en movimiento que se indican como regiones punteadas (fig. 2A). En la figura 2B y 2C se muestra la corrección de distorsiones. (C) Observando la región indicada con la flecha blanca se puede apreciar que se corrigen las curvaturas de distorsión introducidas por la cámara *fisheye*. (D) Sistema de adquisición de imágenes propuesto para este plan compuesto por una cámara *fisheye* y una cámara PTZ.

### 3.b - Modelos estadísticos de las imágenes adquiridas y Clasificación Background-Foreground (BKG-FRG) para la detección de objetos en movimiento:

En los sistemas de VDA se necesitan algoritmos para detectar objetos de interés que se mueven (objetos FRG) respecto del fondo (BKG). Estos algoritmos deben adaptarse a las variaciones de luminosidad del ambiente en distintos momentos del día y también deben poder distinguir por ejemplo, entre el movimiento de ramas de árboles (que corresponden al BKG) respecto de vehículos o personas (que corresponden a objetos FRG de interés). Para segmentar los objetos del FRG se desarrollaron algoritmos heurísticos [7] y algoritmos estadísticos paramétricos y no paramétricos [9].

**3.c - Seguimiento de objetos FRG:** Seguir cada objeto a medida que se mueve por la escena requiere mantener la identidad del objeto a través del tiempo y es claro que esto puede complicarse por problemas de ruido, oclusiones, cruces de objetos, etc. Un enfoque utilizado para resolver estos problemas consiste en filtrar y predecir las trayectorias de los objetos utilizando filtros de Kalman y sus variantes [10]. Un problema que no está completamente resuelto hasta el momento, consiste en mantener la identidad de los objetos entre *frames* sucesivos (que denominamos problema de asignación). En trabajos previos en el contexto del análisis de transporte [3], hemos utilizado los descriptores SURF para tener una descripción detallada de cada punto de interés que facilita la asignación y seguimiento de los objetos.

**3.d - Clasificación y análisis del comportamiento de los objetos FRG:** A medida que se mantiene el seguimiento de los objetos de la escena se pueden estimar para cada objeto distintas propiedades espaciales y temporales. Estas propiedades se utilizan como entradas del cuarto módulo del sistema (Figura 1 (d)) que consiste en un clasificador entrenado por un operador para categorizar los objetos FRG. Las redes

neuronales han demostrado ser buenos clasificadores para estas tareas [11]. También se prevé utilizar técnicas de detección de anomalías [12] que definirán los objetos a enfocar la cámara PTZ para obtener mayor información.

**3.e – Posicionamiento de la cámara PTZ :** Como describimos, los módulos (b-d) determinan los objetos de interés que queremos observar en mayor resolución. Finalmente, el modulo (e) será el encargado controlar el posicionamiento de la cámara PTZ para generar una imagen en alta resolución del objeto de interés.

**Aspectos innovadores del plan de investigación:** *En los últimos años se realizaron avances tanto en el diseño de las cámaras de VCA como en los algoritmos de video-detección automática (VDA) para cámaras convencionales. Ambas líneas de investigación hasta el momento siguieron caminos separados. La innovación principal de este plan consiste en integrar información de campo amplio obtenida a partir de una cámara fisheye con la posibilidad de un análisis en alta resolución controlando el posicionamiento de la cámara PTZ. De este modo, el sistema autónomo tendrá información de todo el campo visual para detectar eventos novedosos e informar a un operador humano.*

**4. Actividades y Metodología:** Se detallan las actividades y metodología prevista para alcanzar los objetivos.

**4.a - Sensores. Sistema de Visión de Campo Amplio (VCA):** En la figura 2D se muestra un esquema del sistema a desarrollar. Para lograr posicionar el centro de la cámara PTZ en una dirección de interés determinada por una posición en la imagen *fisheye*, es necesaria una calibración de los parámetros intrínsecos y extrínsecos de ambas cámaras. Luego de este proceso de calibración quedará definido un mapeo entre la posición de los píxeles de ambas cámaras. En [3] hemos resuelto el problema para la cámara *fisheye*. Por lo tanto nos concentraremos en calibrar los parámetros de posicionamiento y zoom de la cámara PTZ. Luego debemos integrar todo en una descripción georeferenciada a un sistema de referencia fijo al plano de observación. Respecto a la cámara PTZ se debe realizar un estudio detallado del protocolo de comunicación. Finalmente es necesario un análisis de las incertezas de posicionamiento.

**4.b - Clasificación (BKG/FRG) y su extensión a imágenes de VCA:** Este plan prevé analizar la eficiencia de un método heurístico desarrollado por nuestro grupo [2,4] y el modelo paramétrico de *Gaussian Mixture Model* (GMM) propuesto por Stauffer y Grimson [8] y sus variantes. Algunas de las ventajas de este algoritmo son: i) formalismo riguroso desde el punto de vista estadístico y ii) simplicidad del algoritmo con costo computacional relativamente bajo. El algoritmo GMM supone que las clases de píxeles vecinos son independientes y actualmente se está investigando cómo mejorar esta aproximación. En esta dirección, interesa extender el modelo estadístico del GMM y aplicar el método de Máxima Evidencia utilizado previamente en el contexto de clasificación de imágenes médicas [14,15]. Dado que los dos problemas tienen similitudes, se espera realizar contribuciones a la mejora del GMM.

Otro aspecto no solucionado hasta el momento, es el *autoset* de los parámetros de estos algoritmos. La solución de este problema es importante ya que trabajaremos con imágenes en ambientes con iluminación natural que presentan variaciones importantes dependiendo de las condiciones meteorológicas y la posición del sol. Este plan prevé desarrollar un método supervisado que resuelva esta problemática. Para esto un

operador optimizará los parámetros de los algoritmos BKG-FRG en distintas condiciones de adquisición y una red neuronal aprenderá a simular al operador. Finalmente, se evaluará la utilización de técnicas de procesamiento en paralelo para acelerar la aplicación de los algoritmos BKG-FRG.

**4.c - Implementación de algoritmos para seguir objetos de interés:** El método descrito en la sección anterior permite identificar los píxeles FRG en cada nueva imagen adquirida. Estos píxeles FRG deben ser segmentados en regiones conectadas, de modo tal que cada región conectada corresponda a un objeto en movimiento. Para esto, se estudiará un algoritmo tradicional de dos pasadas que demostró ser eficaz en la segmentación los objetos en movimiento [16]. Para realizar el seguimiento de los objetos en movimiento, se propone que se desplazan a velocidad constante. Se estima su estado de posición y velocidad iterativamente utilizando filtros Kalman [10]. Sin embargo, cuando hay distintos objetos en movimiento antes de poder actualizar el estado de cada objeto, es necesario resolver el problema de asignación que mantiene la identidad de los objetos móviles.

El problema de asignación puede pensarse como el problema de minimización de una distancia (o una energía) que depende de los descriptores asociados a cada objeto en los *frames* sucesivos. Por ejemplo en los trabajos [3,4] realizamos algunos avances en el contexto del análisis de imágenes de Sistemas de Transporte. En situaciones relacionadas con la Seguridad el problema presenta mayor complejidad. Por lo que evaluaremos aplicar algoritmos como el presentado en [17].

**4.d - Clasificación y análisis del comportamiento de los objetos FRG:** A medida que se siguen los objetos de la escena, se pueden estimar (para cada uno) propiedades como: tamaño, perímetro, momentos, medidas de dispersión, color, ubicación, velocidad, medidas que cuantifiquen variaciones temporales, etc. Estas propiedades se tomarán como entradas en un clasificador entrenado por un operador para categorizar los objetos FRG en clases como: persona, vehículo, grupo de personas, etc. Las redes neuronales demostraron ser buenos clasificadores para este tipo de tareas [11]. Generalmente se utilizan redes neuronales estándar de tres capas y el aprendizaje en la red se realiza mediante el algoritmo *backpropagation*. También se prevé utilizar técnicas de detección de anomalías [12] que definirán los objetos a enfocar la cámara PTZ para obtener mayor información.

**4.e – Posicionamiento de la cámara PTZ :** En relación a este problema, consideramos evaluar dos estrategias: 1) El sistema VDA+VCA informa la posición de un objeto de interés. La cámara PTZ se ubica para interceptarlo y se establece un lazo de control para el seguimiento del objeto. 2) En función del estado de los objetos en la imagen VCA, se determina una estrategia posicionamiento de la cámara PTZ de manera tal de interceptar el objeto de interés con alta probabilidad, en condiciones favorables y sin la necesidad de seguirlo. La bibliografía informa problemas para el seguimiento de objetos móviles con cámaras PTZ debido a imprecisiones del *timing* de comunicación a través de la red [18]. Creemos que el enfoque 2 producirá un sistema que adquiera imágenes de alta resolución de los objetos de interés en la escena, con mayor eficiencia. Para el control de foco se estudiará la aplicación del método descrito en [13].

## **5- Factibilidad y consideraciones finales:**

**5.1- Factibilidad:** La UNQ brindará las instalaciones necesarias para el desarrollo del proyecto. Las cámaras *fisheye* y PTZ que se necesitan para el desarrollo de este plan ya han sido adquiridas.

*Tareas a desarrollar en el primer año:* Durante el primer año de trabajo, el postulante se concentrará en estudiar detalladamente la bibliografía relacionada con el problema de estudio y en desarrollar el primer objetivo de este problema (módulo 'a'). De alcanzar ambos objetivos, podrá seguir avanzando en este plan.

*Respecto al director y co-director de este proyecto:* Durante el periodo 2013-2014, ambos investigadores trabajaron en esta área de estudio [2-4]. El MSc. Ing. Félix Safar tiene amplia experiencia en investigación y desarrollo de algoritmos y sistemas de procesamiento para visión artificial y procesamiento digital de imágenes, que resultaron en trabajos publicados, patentes de invención, y transferencias en colaboración con *IBM Almaden Research Center*. En este proyecto se concentrará principalmente en los aspectos relacionados con la implementación del sistema *realtime* en hardware y software. El Dr. Damián Oliva es docente de la carrera IACI-UNQ en las materias 'Control de Robots' y 'Redes neuronales y Lógica Difusa', investigador asistente del Conicet en el área de Física. Investiga actualmente temas relacionados con el procesamiento visumotor en sistemas autónomos (biológicos y artificiales). En este proyecto, se concentrará principalmente en los aspectos relacionados con el desarrollo de los algoritmos del sistema.

**Financiamiento:** Este proyecto se enmarcará dentro de los siguientes subsidios:

-Proyecto "Sistemas de Video Detección Vehicular con Visión de Campo Amplio, y su aplicación a

los Sistemas Inteligentes de Transporte (SIT)". Convocatoria "Universidad y Transporte". Ministerio de Educación. 2015. Monto 200000 pesos. Director: Félix Safar, Codirector: Damián Oliva

-Programa de Investigación UNQ 2015-2019. "PUNQ EXPTE 1406/15". Monto 115000 pesos por año. Director: Félix Safar

## **6- Bibliografía**

[1]- <http://www.automationworld.com/computer-vision-opportunity-or-threat>

[2]- Stanganelli D, Oliva D, Noblia M, Safar F. *Calibración de una cámara fisheye comercial con el modelo unificado para la observación de objetos múltiples*. Biennial Congress of Argentina (ARGENCON) 2014 IEEE. DOI:

"<http://dx.doi.org/10.1109/ARGENCON.2014.6868487>" [10.1109/ARGENCON.2014.6868487](https://doi.org/10.1109/ARGENCON.2014.6868487). Pag. 147-152.

[3]- Arroyo S, Safar F, Oliva D. Implementación de un sistema de seguimiento georeferenciado de características en imágenes de campo amplio. Trabajo aceptado en RPIC 2015.

[4]- Oliva D, Yabo G, Garcia L, Arroyo S, Safar F. Implementación de un Sistema para la Medición del Flujo de Tránsito y Detección de Embotellamientos en Autopistas. Trabajo enviado al Simposio Argentino de Inteligencia Artificial. ASAI 2015. Santa Fe, Argentina.

[5]- Hu, W., Tan, T., Wang, L., & Maybank, S. (2004). A survey on visual surveillance of object motion and behaviors. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 34(3), 334-352.

[6]- Fisheye Network Camera FE8172/72V, VIVOTEK INC., 2013.

[7]- Collins, R. T., Lipton, A. J., Fujiyoshi, H., & Kanade, T. (2001). Algorithms for cooperative multisensor surveillance. *Proceedings of the IEEE*, 89(10), 1456-1477.

[8]- Stauffer, C., & Grimson, W. E. L. (1999). Adaptive background mixture models for real-time tracking. In *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on*. (Vol. 2). IEEE.

[9]- McHugh, J. M., Konrad, J., Saligrama, V., & Jodoin, P. M. (2009). Foreground-adaptive background subtraction. *Signal Processing Letters, IEEE*, 16(5), 390-393.

[10]- Black, J., & Ellis, T. (2006). Multi camera image tracking. *Image and Vision Computing*, 24(11), 1256-1267.

- [11]-Lipton, A. J., Fujiyoshi, H., & Patil, R. S. (1998, October). Moving target classification and tracking from real-time video. In *Applications of Computer Vision, 1998. WACV'98. Proceedings., Fourth IEEE Workshop on* (pp. 8-14). IEEE.
- [12]- Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys (CSUR)*, 41(3), 15.
- [13]-Geusebroek, J. M., Cornelissen, F., Smeulders, A. W., & Geerts, H. (2000). Robust autofocusing in microscopy. *Cytometry*, 39(1), 1-9.
- [14]-Oliva, D. E., Isoardi, R. A., & Mato, G. (2008, October). Bayesian estimation of Hyperparameters in MRI through the Maximum Evidence Method. In *Computer Graphics and Image Processing, 2008. SIBGRAPI'08. XXI Brazilian Symposium on* (pp. 129-136). IEEE.
- [15]-Isoardi, R. A., Oliva, D. E., & Mato, G. (2011). Maximum Evidence Method for classification of brain tissues in MRI. *Pattern Recognition Letters*, 32(1), 12-18.
- [16]- Horn B. K. P. *Robot Vision*, pp. 66-69, 299-333. The MIT Press, 1986.
- [17]- Shu, G., Dehghan, A., Oreifej, O., Hand, E., & Shah, M. (2012, June). Part-based multiple-person tracking with partial occlusion handling. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on* (pp. 1815-1821). IEEE.
- [18]- Dinh, T., Yu, Q., & Medioni, G. (2009, October). Real time tracking using an active pan-tilt-zoom network camera. In *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on* (pp. 3786-3793). IEEE.
- [19]- Baricco, Guillermo; Olivero, Alfredo; Rodríguez, Eduardo J.; Safar, Félix G.; Sanz, Jorge (1995). Conformal mapping-based image processing: theory and applications. *Journal of Visual Communication and Image Representation*. Volume 6, Issue 1, March 1995, 35-51.

## **10. TRABAJOS DE INVESTIGACION REALIZADOS O PUBLICADOS EN ESTE PERIODO.**

**10.1 PUBLICACIONES.** *Debe hacer referencia exclusivamente a aquellas publicaciones en la cual se haya hecho explícita mención de su calidad de Becario de la CIC (Ver instructivo para la publicación de trabajos, comunicaciones, tesis, etc.). Toda publicación donde no figure dicha mención no debe ser adjuntada ya que no será tomada en consideración. A cada trabajo asignarle un número e indicar el nombre de los autores, en el mismo orden en que aparecen en la publicación, informe o memoria técnica, lugar donde fue publicado, volumen, página y año si corresponde. En cada trabajo que el becario presente -si lo considerase de importancia- agregará una nota justificando el mismo y su grado de participación. Asimismo, en cada caso deberá indicar si el trabajo se encuentra depositado en el repositorio institucional CIC-Digital.*

**10.2 TRABAJOS EN PRENSA Y/O ACEPTADOS PARA SU PUBLICACIÓN.** *Debe hacer referencia exclusivamente a aquellos trabajos en los que haya hecho explícita mención de su calidad de Becario de la CIC (Ver instructivo para la publicación de trabajos, comunicaciones, tesis, etc.). Todo trabajo donde no figure dicha mención no debe ser adjuntado porque no será tomado en consideración. A cada trabajo, asignarle un número e indicar el nombre de los autores en el mismo orden en que aparecen en la publicación y el lugar donde será publicado. A continuación, transcribir el resumen (abstract) tal como aparecerá en la publicación. La versión completa de cada trabajo se presentará en papel, por separado, juntamente con la constancia de aceptación. En cada trabajo, el becario deberá aclarar el tipo o grado de participación que le cupo en el desarrollo del mismo y, para aquellos en los que considere que ha hecho una contribución de importancia, deberá escribir una breve justificación.*

**10.3 TRABAJOS ENVIADOS Y AUN NO ACEPTADOS PARA SU PUBLICACION.** *Incluir un resumen de no más de 200 palabras de cada trabajo, indicando el lugar al que ha sido enviado. Adjuntar copia de los manuscritos.*

**10.4 TRABAJOS TERMINADOS Y AUN NO ENVIADOS PARA SU PUBLICACION.** *Incluir un resumen de no más de 200 palabras de cada trabajo.*

**10.5 COMUNICACIONES.** *Incluir únicamente un listado y acompañar copia en papel de cada una. (No consignar los trabajos anotados en los subtítulos anteriores).*

**10.6 INFORMES Y MEMORIAS TECNICAS.** *Incluir un listado y acompañar copia en papel de cada uno o referencia de la labor y del lugar de consulta cuando corresponda. Indicar en cada caso si se encuentra depositado en el repositorio institucional CIC-Digital.*

**11. PUBLICACIONES Y DESARROLLOS EN:**

**11.1 DOCENCIA**

**11.2 DIVULGACIÓN**

**11.3 OTROS**

En cada caso indicar si se encuentran depositados en el repositorio institucional CIC-Digital.

**12. PARTICIPACION EN REUNIONES CIENTIFICAS.** *Indicar la denominación, lugar y fecha de realización, tipo de participación que le cupo, títulos de los trabajos o comunicaciones presentadas y autores de los mismos.*

**13. CURSOS DE PERFECCIONAMIENTO, VIAJES DE ESTUDIO, ETC.** *Señalar características del curso o motivo del viaje, período, instituciones visitadas, etc, y si se realizó algún entrenamiento.*

*Machine Learning Course, Buenos Aires 24 – 25 de Noviembre 2016, a cargo de Raúl Ramos Pollan de la Universidad Industrial de Santander, Bucaramanga Colombia. Curso introductorio y de capacitación en Machine learning: Deep learning, Convolutional neural networks, Recurrent networks, Eigenfaces, Autoencoders, Large scale learning, Big data, distributed Machine learning, Machine learning on GPU. El Curso se realizó en el Centro Atómico Constituyentes de la CNEA.*

**14. SUBSIDIOS RECIBIDOS EN EL PERIODO.** *Indicar institución otorgante, fines de los mismos y montos recibidos.*

**15. DISTINCIONES O PREMIOS OBTENIDOS EN EL PERIODO.**

**16. TAREAS DOCENTES DESARROLLADAS EN EL PERIODO.** *Indicar el porcentaje aproximado de su tiempo que le han demandado.*

*Durante el periodo se desarrolló la tarea de instructor contratado con cargo interino en el área de Matemática, puntualmente sobre Análisis 2 con un porcentaje aproximado de tiempo de 16%.*

**17. OTROS ELEMENTOS DE JUICIO NO CONTEMPLADOS EN LOS TITULOS ANTERIORES.** *Bajo este punto se indicará todo lo que se considere de interés para la evaluación de la tarea cumplida en el período.*

**18. DESCRIPCION DEL AVANCE EN LA CARRERA DE DOCTORADO.**

*Debe indicarse los logros alcanzados en la carrera de Doctorado en relación a los requisitos particulares de la misma (cursos, seminarios, trabajos de campo, etc), así como el porcentaje estimado de avance en la tesis.*

**19. TITULO Y PLAN DE TRABAJO A REALIZAR EN EL PROXIMO PERIODO.** *Deberán indicarse claramente las acciones a desarrollar.*

En el próximo periodo se desarrollará un plan de doctorado dentro del contexto de una Beca Doctoral CONICET, cuyo otorgamiento nos fuera informado en Diciembre de 2016. La orientación propuesta es trabajar y desarrollar conocimientos sobre sensores, cámaras, visión artificial, *machine learning* y georeferenciación para su aplicación tanto en ámbitos rurales (agricultura de precisión) como en el análisis automático de escenas urbanas.

**Título: Automatización de procesos basada en video-detección utilizando cámaras móviles en ambientes exteriores.**

.....  
Firma del Director

.....  
Firma del Becario

---

**Condiciones de Presentación**

- A. El Informe Científico deberá presentarse dentro de una carpeta, con la documentación abrochada y en cuyo rótulo figure el Apellido y Nombre del Becario, la que deberá incluir:
- Una copia en papel A-4 (puntos 1 al 14).
  - Las copias de publicaciones y toda otra documentación respaldatoria, deben agregarse al término del desarrollo del informe
  - Informe del Director de tareas con la opinión del desarrollo del becario (en sobre cerrado).

---

**Nota:** El Becario que desee ser considerado a los fines de una prórroga, deberá solicitarlo en el formulario correspondiente, en los períodos que se establezcan en los cronogramas anuales.