

Segmentación de patentes vehiculares mediante técnicas de agrupamiento en ambientes externos

Segmentation of license plates through clustering techniques in external environments

Juan Rojas Henríquez¹ Diego Aracena Pizarro¹

Recibido 26 de noviembre de 2012, aceptado 29 de abril de 2013

Received: November 26, 2012 Accepted: April 29, 2013

RESUMEN

En este trabajo se presenta el empleo de técnicas de agrupamiento para segmentar patentes vehiculares mediante un método aplicable a sistemas ANPR (Automatic Number Plates Recognition). Para lograr este objetivo se realizaron diferentes pruebas con *clusters*, comparando su desempeño y efectividad. Este método utiliza técnicas de agrupamiento para eliminar, de manera automática, los puntos espurios que pueden ser generados por las técnicas de obtención de correspondencias. Para realizar la segmentación de patentes vehiculares, el método recibe como entrada una nube de puntos correspondientes, resultantes de un *matching* entre los descriptores SIFT (Scale Invariant Features Transforms) de la imagen de entrada y los descriptores SIFT de una imagen de referencia que contiene los posibles caracteres de una patente vehicular. Mediante técnicas de agrupamiento se obtiene el *cluster* conformado por los puntos ubicados en la zona de la placa y además su centroide asociado. Luego se emplea la técnica de crecimiento de regiones para segmentar el área mínima que encapsula al centroide, la cual constituye la placa patente de la imagen considerada. En conclusión se presenta un proceso de agrupamiento que forma parte de un método híbrido que combina técnicas de visión e inteligencia computacional para segmentar patentes vehiculares en ambientes externos.

Palabras clave: Segmentación de patentes vehiculares, ANPR (Automatic Number Plates Recognition), técnicas de agrupamiento, técnicas de obtención de correspondencias, SIFT (Scale Invariant Features Transforms), crecimiento de regiones.

ABSTRACT

This paper presents the use of clustering techniques to segment license plates through a method applicable to systems ANPR (Automatic Number Plates Recognition). To achieve this goal were different tests with clusters, comparing its performance and effectiveness. This method uses clustering techniques to remove automatically, spurious points that can be generated by the correlation techniques. To perform segmentation of license plates, the method receives as input a cloud of points corresponding, resulting in a matching between the input image descriptors SIFT (Scale Invariant Features Transforms) and the reference image descriptors SIFT that contains the possible characters of a license plates. Through clustering techniques, gets the cluster consisting of the points located in the zone of the plate and also his associate centroid. Regions growing technique is then used to segment the minimum area that encapsulates the centroid which constitutes the patent plate of the considered image. In conclusion, we present a clustering process that is part of a hybrid approach that combines techniques from computational vision and intelligence, to segment license plates in external environments.

Keywords: Segmentation of license plates, ANPR (Automatic Number Plates Recognition), clustering techniques, getting correspondence techniques, SIFT (Scale Invariant Features Transforms), regions growing.

¹ Área de Ingeniería en Computación e Informática. Universidad de Tarapacá. Av. 18 de Septiembre 2222. Arica, Chile.
E-mail: juanrojashenriquez@gmail.com; daracena@uta.cl

INTRODUCCIÓN

Dentro de los países en desarrollo, el número de vehículos está aumentando día a día. Según las estadísticas de la consultora estadounidense Wards Auto [1], “En 2010 el número de vehículos en circulación en el mundo superó los mil millones”; lo cual significa que en la actualidad hay un vehículo cada siete personas. Este aumento significativo a nivel mundial ha traído consigo la necesidad de contar con sistemas computacionales eficientes para ser utilizados como una herramienta de apoyo a la seguridad y el control de acceso y/o flujo vehicular en distintos escenarios. Entre estos sistemas destacan aquellos orientados al reconocimiento automático de patentes vehiculares, conocidos con el nombre de ANPR (Automatic Number Plates Recognition).

Los sistemas ANPR llevan más de 20 años causando gran interés comercial alrededor del mundo; teniendo diversos campos de aplicación relacionados con la seguridad y el tráfico vehicular. Dentro de estos sistemas, la etapa de detección es clave, puesto que una correcta localización y segmentación de la placa, constituye el primer paso para lograr un buen reconocimiento posterior.

Debido a la importancia que tiene una buena detección y segmentación de la placa, los métodos utilizados deben ser cada vez más eficientes y robustos para enfrentar al máximo los niveles de entropía del escenario en donde se implementará la aplicación. De esta forma es de vital importancia contar con procesos de detección invariantes a la rotación, escala, traslación y luminosidad en ambientes no preparados.

Los métodos usados en la literatura referencial para localizar la placa dentro de la imagen inicial son variados. En el trabajo realizado por F. Martín y X. Fernández [2], se hace uso de operadores morfológicos, en el cual los caracteres son objetos formados por trazos colocados sobre un fondo de color opuesto. Esto permite detectarlos con la transformación morfológica “Top-hat”, que se basa en borrar los caracteres de la imagen original mediante un proceso de “closing” para posteriormente realizar una resta entre ambas ($\text{close}(I)-I$), obteniendo una imagen en la cual aparecerán los caracteres resaltados. Por otro lado, el método de L. Carrera y M. Mora [3] combina operadores morfológicos

con Redes Neuronales Supervisadas (MLP: Multi-Layer Perceptron) para clasificar los descriptores asociados a un conjunto de regiones de interés y seleccionar finalmente la región perteneciente a la placa vehicular.

En el trabajo de S. Rovetta y R. Zunino [4] se utiliza la cuantificación de vectores (Vector Quantization o VQ). Esta técnica se usa, generalmente, para la compresión de datos. Consiste en utilizar un libro de código predefinido en el que cada entrada lleva asociada una probabilidad de formar parte de la placa. De esta forma se busca aumentar la posibilidad de localizar con precisión el área donde se encuentra la patente vehicular.

El método propuesto por F. Wang, L. Man, B. Wang, Y. Xiao, W. Pan y X. Lu [5] se centra en el espacio de color (HSV) de la placa para la extracción de sus características mediante la aplicación de lógica difusa (Fuzzy Logic). Aquí la imagen del vehículo es particionada en rectángulos de un tamaño aproximado al ocupado por la placa. En estos rectángulos se calcula la membresía de conjuntos difusos, creados a partir de reglas intuitivas como el área aproximada de localización de la placa, la brillantez del borde, etc. Mientras que en el trabajo de V. Abolghasemi y A. Ahmadyfard [6], se emplea la técnica MNS (Multimodal Neighborhood Signature) para utilizar la textura del color de la placa como una señal para su detección.

En el método propuesto por J. Jiao, Q. Ye y Q. Huang [7] se presenta un esquema configurable de reconocimiento de placas multiestilo, el cual consiste en la adaptación de parámetros como: ángulo de rotación, número de líneas, tipos de caracteres y formatos de caracteres, de acuerdo con las características de las placas. Esta cualidad permite obtener el reconocimiento de placas de distintas nacionalidades.

En el trabajo realizado por J. Rojas y D. Aracena [8] se presenta un método basado en gradientes para localizar y segmentar la placa vehicular. Inicialmente se detectan los bordes verticales de la imagen, mediante el operador de derivación de Sobel. Posteriormente se calcula un umbral para clasificar la imagen gradiente en tres posibles valores (-1, 0, 1), que junto con parámetros inicialmente conocidos de la placa, como sus dimensiones,

distancia entre caracteres y ancho de sus trazos, se utilizan para segmentar la placa vehicular mediante las características de la firma. Finalmente, la imagen resultante se binariza y se extrae todo el “ruido” que presenta para obtener únicamente los caracteres de la placa vehicular.

Por último, en el trabajo de M. Zahedi y S. Mahdi [9] se hace uso de SIFT [10] para el reconocimiento de placas vehiculares conocidas. Este método consiste en detectar los puntos característicos y los respectivos descriptores de una imagen. Con la información proporcionada por SIFT se realizan distintos *matching's* entre los descriptores de la imagen de entrada y los descriptores de distintas imágenes de placas vehiculares almacenadas en el sistema (las cuales muestran la totalidad de sus caracteres). Este proceso se realiza con la finalidad de encontrar el *matching* que entregue el mayor número de correspondencias entre la imagen de entrada y alguna de las placas vehiculares almacenadas en el sistema. Es decir, se realizan tantos *matching's* como tantas placas vehiculares se encuentren disponibles en la base de datos, por esta razón, el método de M. Zahedi y S. Mahdi [9] está orientado a escenarios específicos en los cuales se conocen las placas de los vehículos que están autorizados para ingresar a un determinado recinto y que serán finalmente aquellas placas que estarán almacenadas en la base de datos del sistema. Para aumentar la efectividad del método de M. Zahedi y S. Mahdi [9], los autores recomiendan realizar una etapa previa antes de la aplicación de SIFT y los distintos *matching's*. Esta etapa consiste en procesar la imagen de entrada utilizando un operador de detección de bordes verticales para luego segmentar la patente vehicular mediante un método basado en gradientes. De esta forma, la imagen de entrada se reducirá a una imagen que mostrará únicamente la placa vehicular, lo cual disminuirá considerablemente el número de cálculos involucrados en los *matching's* y aumentará el porcentaje de éxito al momento de reconocer la placa vehicular. Sin embargo, debido a la cantidad de *matching's* que debe realizar el sistema, el tiempo de ejecución empleado para realizar este proceso resultará ser demasiado alto, ya que depende directamente de la cantidad de placas vehiculares que se encuentren almacenadas. Además, la etapa previa que recomiendan los autores, está basada en técnicas de detección de bordes y métodos basados en gradientes que funcionan únicamente en ambientes

controlados en los que los parámetros asociados a la placa, como sus dimensiones, distancia entre caracteres, ancho de sus trazos y distancia entre la cámara y el vehículo, siempre deben cumplirse, ya que cualquier variación en alguno de estos parámetros, ocasionaría una segmentación errónea. Debido a estas razones, el método de M. Zahedi y S. Mahdi [9] sería inviable para ser aplicado a ambientes externos, no controlados y en los cuales no se conozcan las patentes vehiculares que pueden ingresar al sistema.

Este trabajo, inspirado en [9], presenta una arquitectura para segmentar placas patentes, no conocidas, mediante técnicas de agrupamiento en ambientes externos. Para lograr este cometido se utiliza SIFT para detectar los puntos característicos y los descriptores de una imagen de entrada y de una imagen de referencia, conformada por caracteres extraídos de distintas imágenes de prueba, que corresponden a letras y números, normados, de distintos tamaños, pertenecientes a placas de diferentes estilos. Luego, se realiza un único *matching* para obtener los puntos correspondientes entre ambas imágenes y posteriormente, como parte fundamental de este proceso, se aplican las técnicas de agrupamiento *K-means* [11], *Fuzzy c-means* [12] y *Mean shift* [13], para determinar el *cluster* que representa a la placa vehicular. Finalmente se aplica la técnica de crecimiento de regiones para segmentar la placa patente que deberá pasar a la etapa de reconocimiento de caracteres para finalizar el proceso. La principal contribución de este método radica en la forma en cómo se realiza la segmentación de la placa vehicular, ya que se trata de un método híbrido que combina técnicas de visión computacional e inteligencia computacional. La utilización del detector SIFT, junto con técnicas como el *matching*, agrupamiento (*clustering*) y crecimiento de regiones hacen posible que el método propuesto sea capaz de detectar y segmentar placas vehiculares no conocidas, independientemente del estilo que estas posean y sin tener la necesidad de conocer *a priori*, parámetros propios de la placa vehicular que se requiera detectar. Este método responde a la necesidad de contar con procesos de detección invariantes a la rotación, escala, traslación y luminosidad en ambientes no preparados. Esta característica hace que este método pueda ser aplicado en sistemas ANPR orientados al control de acceso y/o flujo vehicular en ambientes externos.

ARQUITECTURA DEL SISTEMA

Este trabajo se enfoca en la primera etapa de un sistema ANPR, la cual consiste en detectar y segmentar la placa que contiene los caracteres a reconocer. Esta etapa es fundamental para el sistema, puesto que solo si se obtiene una correcta detección y segmentación de la placa se podría llegar a alcanzar un exitoso reconocimiento posterior.

En la Figura 1 se muestra la arquitectura del sistema propuesto, la cual está compuesta por cuatro etapas principales:

- SIFT (puntos característicos y descriptores)
- Matching entre descriptores SIFT
- Técnicas de Agrupamiento (*Clustering*)
- Crecimiento de regiones

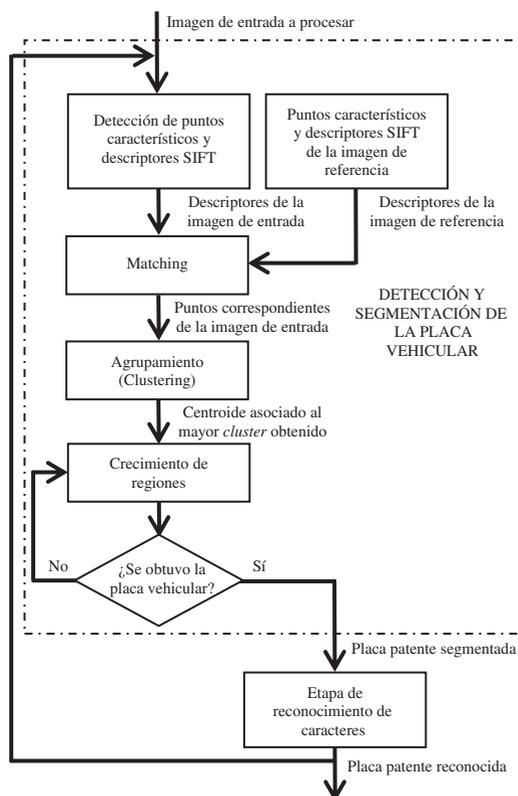


Figura 1. Arquitectura del sistema de segmentación de patentes vehiculares basado en SIFT y técnicas de agrupamiento en ambientes externos.

METODOLOGÍA

SIFT: Detección invariante de puntos característicos y descriptores de la imagen

Uno de los métodos de detección de puntos invariantes ampliamente utilizados en la literatura es SIFT (Scale Invariant Features Transforms) [10]. Este detector transforma los datos de la imagen en descriptores invariantes a la traslación, escala, rotación y en cierta medida al cambio de iluminación. Cada uno de los puntos extraídos se considera una característica de la imagen y se describe mediante su posición, escala, orientación y su vector descriptivo, el cual contiene 128 posibles atributos (bins de histogramas).

El algoritmo SIFT se compone principalmente de cuatro etapas para extraer las características de la imagen:

- Detección de máximos y mínimos en el espacio escala
- Localización y filtrado de puntos característicos
- Asignación de la orientación
- Generación del descriptor del punto característico

SIFT es un detector afín que logra identificar un gran número de características, lo cual es particularmente importante para la detección de la placa vehicular dentro de la imagen. La alta distintividad de los descriptores que genera SIFT otorga la posibilidad de poder comparar puntos pertenecientes a un mismo objeto ubicado en distintas imágenes. Estos reciben el nombre de puntos correspondientes, los cuales se determinan utilizando técnicas de obtención de correspondencias (*matching*).

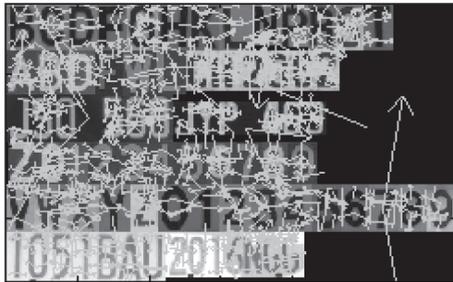
En este trabajo se utiliza SIFT para extraer los puntos característicos, junto con sus respectivos descriptores, de la imagen de entrada a procesar que corresponde a un vehículo y de la imagen de referencia que contiene los posibles caracteres de una patente vehicular.

En una etapa previa a la ejecución del método propuesto se almacenan los puntos característicos y descriptores SIFT de la imagen de referencia. Estos caracteres, extraídos de distintas imágenes de prueba, corresponden a letras y números, normados, de distintos tamaños pertenecientes a placas de diferentes estilos. Los descriptores SIFT asociados a esta imagen de referencia serán utilizados

posteriormente en el proceso de *matching*, en forma conjunta, con los descriptores de la imagen de entrada. En la Figura 2 se muestran los puntos característicos y los descriptores asociados a una posible imagen de entrada y a la imagen de referencia.



(a)



(b)

Figura 2. SIFT: Puntos característicos y descriptores, (a) posible imagen de entrada (630 pts.) y (b) imagen de referencia (1016 pts.).

Matching entre descriptores SIFT

Una vez que se obtienen los descriptores asociados a los puntos característicos de la imagen de entrada y la imagen de referencia, se realiza un *matching* basado en el vecino más cercano con relación de distancias entre el primer y segundo vecino más cercano, con la finalidad de obtener los puntos correspondientes entre ambas imágenes. Para calcular el grado de similitud entre dos puntos característicos se utiliza distancia euclidiana, la cual se muestra en la ecuación (1).

$$d_{total} = \sqrt{\sum_{k=1}^n (a_i - b_i)^2} \quad (1)$$

En donde:

- $k=1...n$ (n representa el número de descriptores de cada punto característico, en este caso $n = 128$).

- Para ambos puntos característicos en proceso se realiza una diferencia euclidiana entre el descriptor a_i de la imagen de entrada y el descriptor b_i de la imagen de referencia que contiene los posibles caracteres de una patente vehicular.
- d_{total} es la suma de las distancias euclidianas que hay entre los 128 descriptores de ambos puntos característicos en proceso.
- Se elige siempre la mínima d_{total} obtenida, puesto que esto sería un indicio de que entre estos dos pares de puntos en cuestión existe un alto grado de correspondencia.
- Finalmente la cantidad de puntos obtenidos se filtran mediante un umbral, por lo que solamente un porcentaje de estos formarán parte de los puntos correspondientes definitivos.

Luego de realizar el *matching* entre ambas imágenes, como se muestra en la Figura 3, se utiliza únicamente la nube de puntos correspondientes de la imagen de entrada, los cuales, en su mayoría, se encuentran ubicados en el interior de la placa vehicular, pero también habrá puntos espurios situados en otros lugares de la imagen que no proporcionarán información relevante, por lo que deben ser eliminados de manera automática.



Figura 3. Puntos correspondientes entre la imagen de entrada y la imagen de referencia. (umbral = 0,80, puntos correspondientes = 39).

Técnicas de Agrupamiento (*Clustering*)

En esta etapa se aplican técnicas de agrupamiento para excluir, de forma automática, aquellos puntos correspondientes “espurios” obtenidos en el proceso de *matching*.

Las técnicas de agrupamiento, también conocidas como clasificación no supervisada, son un método para generar grupos de objetos (*clusters*), de tal manera que aquellos elementos pertenecientes a

un mismo *cluster* tengan un alto grado de similitud entre ellos, y a la vez, un bajo grado de similitud con elementos pertenecientes a otros *clusters*.

El objetivo de esta etapa es localizar el *cluster* conformado por los puntos ubicados en la zona de la placa vehicular y además obtener su centroide asociado. Esto se puede lograr debido a que los puntos correspondientes de la imagen de entrada se obtienen utilizando una imagen de referencia; es por esto que, generalmente, el *cluster* que representa a la placa suele ser aquel que contiene la mayor cantidad de puntos correspondientes, tal como se muestra en la Figura 4. Para lograr el objetivo planteado se aplicaron las técnicas *K-means* (agrupamiento clásico), *Fuzzy c-means* (agrupamiento difuso) y *Mean shift* (agrupamiento basado en densidad), las cuales fueron analizadas y comparadas para ver cuál de ellas otorgaba los mejores resultados.

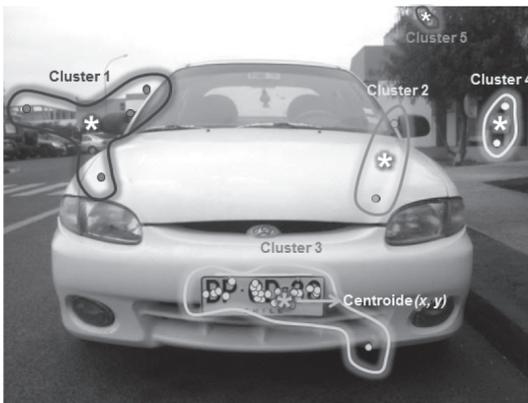


Figura 4. Representación gráfica de agrupamiento utilizando cinco *clusters*.

Inicialmente se tiene un conjunto de elementos para los que se define el número de *clusters* en los que serán agrupados y sus centroides iniciales. Aquellos elementos que estén más cercanos a un centroide en específico formarán parte del *cluster* asociado a ese centroide. Una vez que se conformen los primeros *clusters* se actualizan los centroides siguiendo el mismo criterio de distancia y los elementos se vuelven a distribuir entre los *clusters*, hasta alcanzar una distribución estable.

Estas técnicas se clasifican en paramétricas y no paramétricas. A su vez, los algoritmos de agrupamiento no paramétricos pueden dividirse en

tres grupos fundamentales: jerárquicos, particionales y basados en densidad. En este trabajo se utilizan específicamente técnicas pertenecientes a los dos últimos grupos.

K-means (Agrupamiento clásico)

K-means [11] es uno de los algoritmos de agrupamiento particionales más conocido y utilizado. Este algoritmo realiza la asignación de elementos, dentro de k *clusters* definidos *a priori*, basándose en el criterio de la mínima distancia respecto del centroide de cada grupo, de tal manera que cada uno de los *clusters* queda representado por la media (o media ponderada) de sus puntos, es decir, por su centroide. *K-means* tiene una complejidad de $O(nkl)$, donde n es el número de elementos, k es el número de *clusters* y l es el número total de iteraciones necesarias para alcanzar la convergencia.

El algoritmo de agrupamiento *K-means* funciona tal como se muestra en la Figura 5.

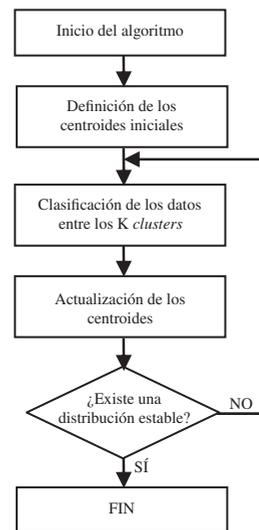


Figura 5. Diagrama de flujo del algoritmo *K-means*.

En la etapa de clasificación se determina a qué *cluster* pertenecerá cada punto. Para esto se calcula la distancia existente entre el punto en cuestión y cada centroide. La pertenencia de cada punto a un determinado *cluster* queda definida por la menor distancia entre el punto y el centroide asociado al *cluster*. Los tipos de distancias que se utilizaron para lograr la clasificación fueron: distancia euclidiana y distancia *City block*. Para esto se utilizan las ecuaciones (2) y (3).

$$d_{ij} = \sqrt{\sum_{k=1}^n (z_{kj} - v_{ki})^2} \quad (2)$$

$$d_{ij} = \sum_{k=1}^n |z_{kj} - v_{ki}| \quad (3)$$

En donde:

- $k=1...n$ (n representa el número de variables para representar un punto, plano $n=2$).
- $i=1...c$ (c representa el número de clusters).
- $j=1...N$ (N representa el número de puntos de cada variable).

En la Figura 6 se muestra la forma en que *K-means* agrupa los puntos correspondientes de la imagen de entrada.

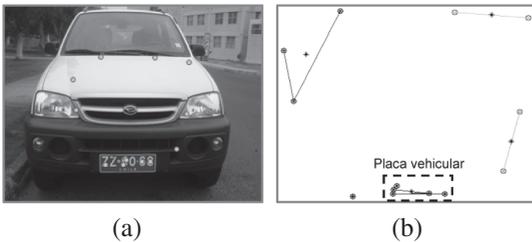


Figura 6. Agrupamiento mediante *K-means*, (a) puntos correspondientes de la imagen de entrada y (b) *K-means* (5 iteraciones, 5 clusters, distancia *City block*).

Fuzzy c-means (Agrupamiento difuso)

La técnica de agrupamiento difuso *Fuzzy c-means* [12] permite conocer el grado con que los elementos son agrupados. Esta ambigüedad (*fuzziness*) refleja el hecho de que un elemento puede pertenecer a más de un grupo. El grado de *fuzziness* de un elemento está dado por un coeficiente, por lo que un valor elevado del mismo significa que el elemento tiene una alta ambigüedad en su asignación. Este algoritmo se basa en la optimización de una función objetivo, la cual se muestra en la ecuación (4).

$$J(U, c_1, \dots, c_c) = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m d_{ij}^2 \quad (4)$$

En donde:

- $i=1...c$ (c representa el número de clusters).
- $j=1...n$ (n representa el número de puntos de cada variable).

- U es la matriz de pertenencias de tamaño $n \times c$. En este sentido, el elemento u_{ki} representa el grado de *fuzziness* que tiene el i -ésimo dato respecto del cluster c_j . Mientras que m es el cociente que controla el grado de pertenencia.
- d_{ij}^2 es la medida de similitud definida, la cual se basa en el criterio del error cuadrático.

El algoritmo de agrupamiento *Fuzzy c-means* funciona tal como se muestra en la Figura 7.

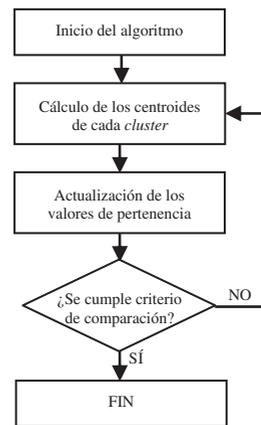


Figura 7. Diagrama de flujo del algoritmo *Fuzzy c-means*.

La forma en que se calculan los centroides de cada cluster se muestra en la ecuación (5).

$$V_i = \frac{\sum_{j=1}^J (\mu_{ij})^m X_j}{\sum_{j=1}^J (\mu_{ij})^m}, \forall i = 1, \dots, c \quad (5)$$

En donde:

- J : número de datos a agrupar.
- $X_j; j=1, \dots, J$ (vector de atributos del dato J).
- c : número de clusters.
- m : grado de *fuzziness* para los clusters encontrados.
- $\mu_{ij}; i=1, \dots, c; j=1, \dots, J$ (grado de pertenencia del dato j al conjunto i).

La actualización de los valores de pertenencia se realiza mediante la ecuación (6).

$$\mu_{ij} = \frac{1}{\sum_{k=1}^c \left[\frac{d_{ij}}{d_{kj}} \right]^{2/m-1}} \quad (6)$$

$$\forall i = 1, \dots, c; \forall j = 1, \dots, J$$

En donde:

- d_{ij} : distancia entre el dato j y el centroide V_i del *cluster* i .

Finalmente, se verifica la condición de convergencia de *Fuzzy c-means*, la cual se muestra en la ecuación (7).

$$\left\| U^{(i)} - U^{(i-1)} \right\| \leq \varepsilon_F \quad (7)$$

En donde:

- $U^{(i)}$ es la matriz de pertenencias de la iteración actual.
- $U^{(i-1)}$ es la matriz de pertenencias de la iteración anterior.
- ε_F es un criterio de particionado para los centroides de los *clusters*.

En la Figura 8 se muestra la forma en que *Fuzzy c-means* agrupa los puntos correspondientes de la imagen de entrada.

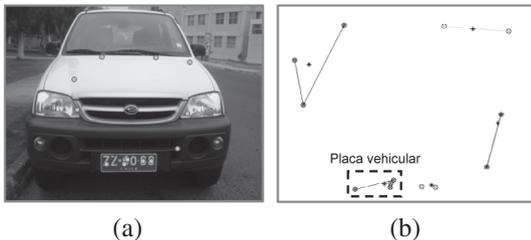


Figura 8. Agrupamiento mediante *Fuzzy c-means*, (a) puntos correspondientes de la imagen de entrada y (b) *Fuzzy c-means* (50 iteraciones máximas, 5 *clusters*, exp. matriz de pertenencia = 2, dif. variaciones centroides = $1e-5$).

Mean shift (Agrupamiento basado en densidad)
Mean shift [13] constituye una técnica de agrupamiento no paramétrica basada en densidad. Esta técnica, en contraste con *K-means* y *Fuzzy c-means*, no requiere conocer *a priori* el número de *clusters* en los que se distribuirán los elementos. Este algoritmo trata el problema de agrupamiento de un espacio d -dimensional como una función de densidad de probabilidad, la cual se muestra en la ecuación (8). En esta función, los grupos corresponden a los máximos y sus puntos máximos son interpretados como el centroide de sus elementos.

Para cada punto de los elementos por agrupar se realiza un procedimiento para calcular el gradiente de la función de densidad hasta su convergencia.

$$\hat{f}(x) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{(x-x_i)}{h}\right) \quad (8)$$

En donde:

- $i=1..n$ (n representa el número de puntos).
- $d=2$ (ya que los puntos pertenecen a un espacio de 2 dimensiones).
- x_1, \dots, x_n son los puntos del conjunto de datos.
- K es el *kernel* para hallar la función de densidad.
- h es el valor de *ancho de banda* a definir.

En la Figura 9 se muestra la forma en que *Mean shift* agrupa los puntos correspondientes de la imagen de entrada.

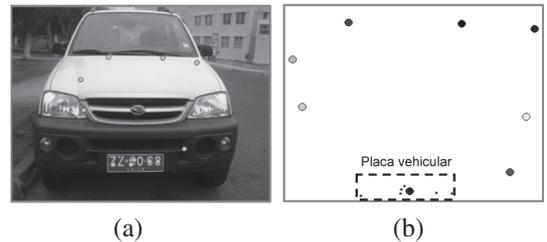


Figura 9. Agrupamiento mediante *Mean shift*, (a) puntos correspondientes de la imagen de entrada y (b) *Mean shift* (ancho de banda = 50).

Crecimiento de regiones

En esta etapa se recibe la imagen de entrada y el centroide asociado a la región que se desea segmentar. El resultado de aplicar esta técnica es una subimagen que corresponde a la menor región que encapsula al centroide.

Cabe señalar que este proceso debe ser controlado mediante un umbral que se adapte a la luminosidad de cada una de las imágenes de entrada, de tal manera que se logre obtener la placa vehicular íntegramente en la imagen binarizada, de esta forma se podrá garantizar una correcta segmentación de la misma. Esta situación se muestra en la Figura 10. Aquí se puede apreciar una misma imagen binarizada con distintos valores de umbral, siendo la Figura 10-c la que muestra la placa vehicular de manera íntegra, lo que la convierte en la imagen candidata para obtener una correcta segmentación de la placa.

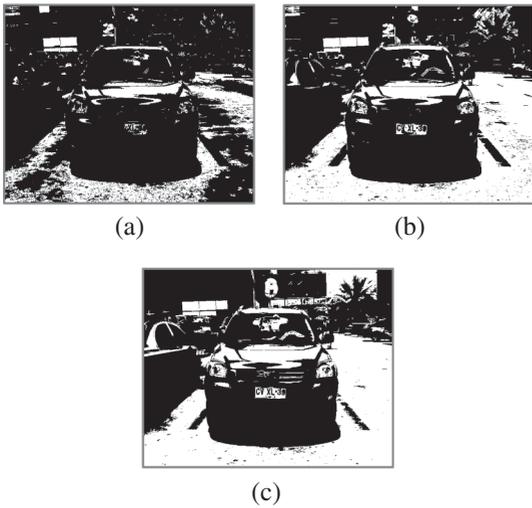


Figura 10. Imagen binarizada mediante umbral (a) umbral = 0,10, (b) umbral = 0,20 y (c) umbral = 0,30 (valor correcto).

Otra situación que puede acontecer en esta etapa es aquella en la cual el centroide se sitúa en el interior de la región de un carácter de la placa patente, esto por lo general origina una segmentación errónea que no corresponde a la placa vehicular. Esta situación puede evitarse, ya que el total de píxeles aproximados que ocupa la placa patente dentro de la imagen inicial no puede ser un valor muy pequeño. De esta manera, el centroide puede ser desplazado hasta quedar situado fuera del carácter y rodeado únicamente por el marco de la placa; situación que genera una segmentación correcta que corresponde justamente a la patente vehicular. Finalmente, la placa segmentada deberá ingresar a la etapa de reconocimiento de caracteres, en la cual se pueden aplicar diversas técnicas, como por ejemplo Redes Neuronales, Lógica Difusa o Reconocimiento Óptico de Caracteres (OCR).

PRUEBAS

Las pruebas fueron realizadas en Matlab®. Se utilizó un computador Intel® Core™2 Duo CPU T6600 de 2.20 GHz de velocidad y 2,00 GB de memoria RAM. Se capturaron imágenes VGA (4:3) y videos (60 Fps) de 640x480 píxeles con dos cámaras digitales: SAMSUNG Digimax A40 de 4.0 megapíxeles y SONY DSC-HX9V de 16.2 megapíxeles.

Para realizar las pruebas se tomaron 120 imágenes, al azar, de vistas frontales y traseras de vehículos chilenos, peruanos, bolivianos y argentinos, ubicados en las calles de la ciudad. Las imágenes se muestran en la Figura 11.

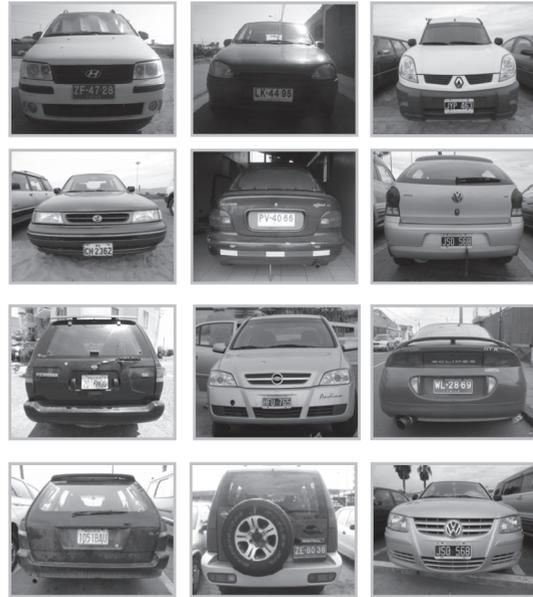


Figura 11. Parte del set de imágenes que se tomaron para realizar las pruebas.

Placas vehiculares de distintos tamaños

La Figura 12 muestra pruebas realizadas con vehículos ubicados a diferentes distancias de la cámara. Estas imágenes permiten probar la respuesta del método propuesto frente a placas vehiculares que presenten distintos tamaños.

Placas vehiculares inclinadas

La Figura 13 muestra pruebas realizadas con vehículos que tienen sus placas inclinadas. Estas imágenes permiten probar la respuesta del método propuesto frente a placas vehiculares que presenten distintos ángulos de inclinación.

Placa vehicular en distintas posiciones de la imagen

La Figura 14 muestra pruebas realizadas con un mismo vehículo, el cual fue capturado por la cámara en distintas ubicaciones. Estas imágenes permiten probar la respuesta del método propuesto frente a la situación en que la placa vehicular se encuentre en distintas posiciones de la imagen.

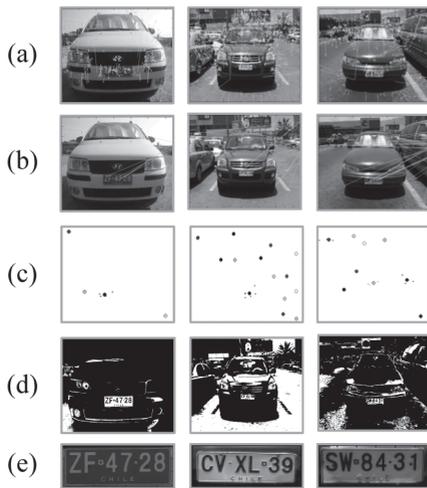


Figura 12. Vehículos ubicados a distintas distancias de la cámara, (a) puntos característicos y descriptores (514 pts., 1726 pts., 1270 pts.), (b) puntos correspondientes de la imagen de entrada (10 pts., 23 pts., 22 pts.), (c) agrupamiento *Mean shift*, (d) binarización mediante umbral (0,10, 0,30, 0,10) y (e) placa vehicular segmentada.

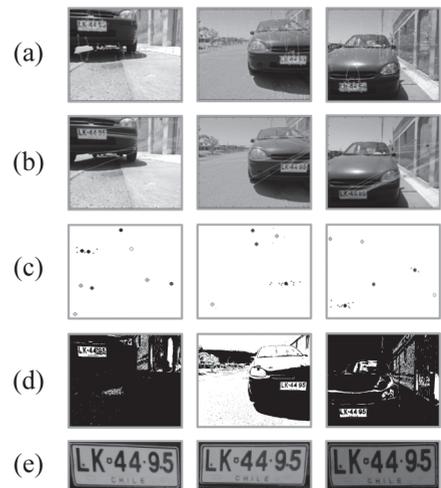


Figura 14. Placa vehicular en distintas posiciones de la imagen, (a) puntos característicos y descriptores (1021 pts., 845 pts., 732 pts.), (b) puntos correspondientes de la imagen de entrada (15 pts., 21 pts., 22 pts.), (c) agrupamiento *Mean shift*, (d) binarización mediante umbral (0,10, 0,40, 0,10) y (e) placa vehicular segmentada.

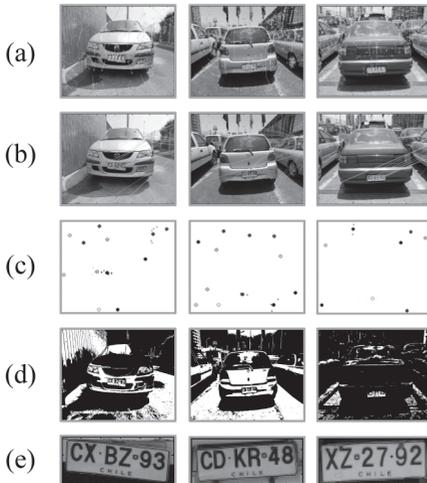


Figura 13. Vehículos que tienen sus placas inclinadas, (a) puntos característicos y descriptores (817 pts., 1380 pts., 1698 pts.), (b) puntos correspondientes de la imagen de entrada (34 pts., 19 pts., 22 pts.), (c) agrupamiento *Mean shift*, (d) binarización mediante umbral (0,20, 0,40, 0,10) y (e) placa vehicular segmentada.

Vehículo con placa inclinada y con un objeto al frente

La Figura 15 muestra un vehículo que tiene su placa inclinada y además tiene un objeto al frente. Este objeto es un letrero de publicidad que contiene varios caracteres, lo cual significa un desafío al momento de realizar el *matching* entre descriptores SIFT, debido a que la correspondencia entre los puntos de la imagen de referencia y los puntos ubicados en la

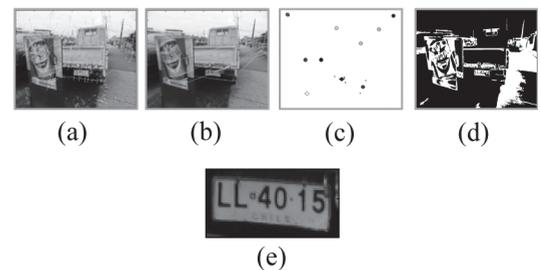


Figura 15. Vehículo con placa inclinada y con un objeto al frente, (a) puntos característicos y descriptores (1063 pts.), (b) puntos correspondientes de la imagen de entrada (18 pts.), (c) agrupamiento *Mean shift*, (d) binarización mediante umbral (0,40) y (e) placa vehicular segmentada.

zona del letrero de publicidad, podría ser mayor con relación a los puntos ubicados en la zona de la placa vehicular. Esto provocaría una segmentación errónea de alguna zona del letrero, en lugar de segmentar la placa vehicular. En esta prueba, el *matching* se comportó de manera esperada, por lo que se pudo segmentar la placa vehicular de manera exitosa.

Placas vehiculares de distintos estilos

La Figura 16 muestra pruebas realizadas con vehículos con placas patentes de distintas nacionalidades (Perú, Bolivia y Argentina). Estas imágenes permiten probar la respuesta del método propuesto frente a placas vehiculares de distintos estilos.

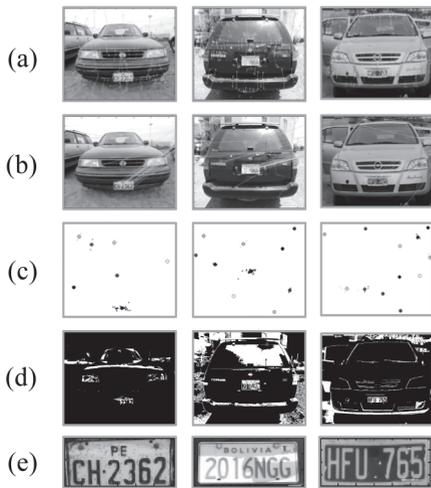


Figura 16. Placas vehiculares de distintos estilos, (a) puntos característicos y descriptores (817 pts., 1302 pts., 955 pts.), (b) puntos correspondientes de la imagen de entrada (42 pts., 43 pts., 24 pts.), (c) agrupamiento *Mean shift*, (d) binarización mediante umbral (0,10, 0,50, 0,20) y (e) placa vehicular segmentada.

Vehículo capturado en un video

La Figura 17 muestra un vehículo capturado en un video grabado a 60 Fps. Para realizar esta prueba se posicionó la cámara, en perspectiva, sobre un trípode y a una distancia aproximada de 4 m. Es importante señalar que la velocidad de grabación del video se fijó en 60 Fps. para garantizar la obtención de un *frame* de calidad, puesto que el vehículo se encontraba en movimiento al momento de la captura.

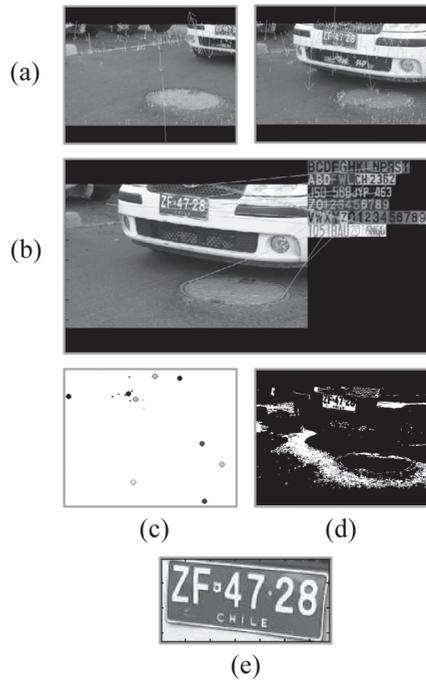


Figura 17. Vehículo capturado en un video (a) puntos característicos y descriptores del *frame* actual (454 pts.) y *frame* siguiente (940 pts.), (b) puntos correspondientes entre la imagen de entrada y la imagen de referencia (19 pts.), (c) agrupamiento *Mean shift*, (d) binarización mediante umbral (0,90) y (e) placa vehicular segmentada.

RESULTADOS

La Tabla 1 muestra una comparación del número de detecciones exitosas que se obtuvo utilizando la técnica de agrupamiento *K-means* (5 iteraciones), empleando diferentes medidas de distancia (euclidiana y *City block*) y número de *clusters* (2 y 5).

Tabla 1. Número de detecciones exitosas con *K-means* (5 iteraciones) utilizando diferentes medidas de distancia y número de *clusters*.

Distancia	N° de aciertos	
	2 clusters	5 clusters
Euclidiana	55	81
City block	87	90

La Tabla 2 muestra una comparación del número de detecciones exitosas y los tiempos de ejecución obtenidos en Matlab®, infiriendo de acuerdo con la experiencia, el tiempo de ejecución en milisegundos para el caso compilado del proceso de detección de la placa patente, según la técnica de agrupamiento utilizada.

Tabla 2. Número de detecciones exitosas y tiempos de ejecución, en milisegundos, del proceso de detección de la placa patente, según la técnica de agrupamiento utilizada.

Agrupamiento	Nº de aciertos	Tiempos en milisegundos		
		mín.	máx.	promedio
<i>K-means</i> ^a	90	21	36	30
<i>Fuzzy c-means</i> ^b	93	21	39	29
<i>Mean shift</i> ^c	97	21	38	28

- a. 5 iteraciones, 5 clusters, distancia *City block*.
- b. 50 iteraciones máx., 5 clusters, exp. matriz de pertenencia = 2, dif. variaciones centroides = 1e-5.
- c. ancho de banda = 50.

De las tres técnicas de agrupamiento que fueron utilizadas en el proceso de detección de la placa vehicular, *Mean shift*^c fue la que proporcionó los mejores resultados en cuanto a número de detecciones exitosas y tiempo de ejecución. Con esta técnica se lograron detectar 97 imágenes de las 120 disponibles, obteniendo así un 80,83% de porcentaje de acierto en la detección de la placa patente. Mientras que las técnicas *K-means*^a y *Fuzzy c-means*^b obtuvieron un porcentaje de acierto del 75% y 77,5% respectivamente.

El método de M. Zahedi y S. Mahdi [9] alcanza un porcentaje de acierto del 84%, pero es necesario recalcar que se trata de un método que funciona en base a placas vehiculares conocidas y que está sujeto a técnicas que funcionan únicamente en ambientes controlados en los que parámetros propios de la placa vehicular siempre deben cumplirse, ya que cualquier variación en alguno de estos ocasionaría una segmentación errónea. Además, su principal limitante es el elevado tiempo de ejecución que necesita para realizar la totalidad de los *matching*'s requeridos, lo que lo hace inviable para ser aplicado en escenarios externos.

El método propuesto en este artículo alcanza un 81% de porcentaje de acierto. Cabe señalar que este método no requiere conocer las placas de los vehículos que puedan conformar las imágenes de entrada y está orientado a funcionar en ambientes externos, no controlados, que no están sujetos a ningún tipo de parámetro asociado a la placa vehicular. Además, no cuenta con ninguna limitante relacionada con el tiempo de ejecución. Por esta razón, los resultados logrados en este trabajo superan en diversos aspectos a los obtenidos por el método de M. Zahedi y S. Mahdi [9].

Cabe destacar que la conversión del código a lenguaje C/C++ reduciría significativamente los tiempos de ejecución obtenidos al orden de los milisegundos, lo cual es clave para una futura implementación, puesto que el tiempo es una variable crítica en este tipo de sistemas.

CONCLUSIONES

La utilización de SIFT para detectar los puntos característicos de la imagen de entrada y la imagen de referencia que contiene los posibles caracteres de una placa patente, es fundamental para el proceso de detección en ambientes externos, puesto que se trata de un detector que proporciona invariancia a ciertos aspectos de la placa (rotación, escala, traslación y en cierta medida al cambio de iluminación) y además genera los descriptores necesarios para realizar el proceso de *matching*.

La aplicación de técnicas de agrupamiento es primordial para eliminar de forma automática los puntos correspondientes espurios que entrega el proceso de *matching*. Estas técnicas permiten obtener el *cluster* que representa a la placa y su centroide asociado, el cual se utiliza para segmentar la placa vehicular mediante la técnica de crecimiento de regiones.

Respecto de las distancias utilizadas en *K-means*, euclidiana responde de mejor manera a mayor cantidad de *clusters*. Mientras que la distancia *City block* resulta ser la más estable y la que mejores resultados entrega, independientemente del número de *clusters* empleados. Esto se debe a que tanto euclidiana como *City block* tienen un concepto

distinto de distancia. Para euclidiana, la distancia mínima entre dos puntos es la recta directa que los une, mientras que *City block* se mueve a lo largo de los ejes horizontales y verticales. De esta forma, para un mismo par de puntos, se van a obtener distintos valores de distancias, siendo *City block* la que mejores resultados entregó.

Finalmente se puede decir que este método logra detectar placas patentes en ambientes externos de manera satisfactoria y en un bajo tiempo de ejecución.

TRABAJO FUTURO

En una primera instancia se continuarán investigando nuevas técnicas de agrupamiento y se compararán con las ya consideradas, de manera de adoptar la técnica que permita obtener los mejores resultados, con la finalidad de poder aplicarla en sistemas orientados al control de acceso y/o flujo vehicular. Además, se aplicarán técnicas de reconocimiento para poder validar la placa patente segmentada y obtener información sobre ella mediante una base de datos.

REFERENCIAS

- [1] J. Sousanis. "World Vehicle Population Tops 1 Billion Units". Agosto 2011. Fecha de consulta: 01 de Mayo 2012. URL: http://wardsauto.com/ar/world_vehicle_population_110815.
- [2] F. Martín y X. Fernández. "Un nuevo método, basado en morfología, para la localización de matrículas". 15° Simposio Nacional de la Unión Científica Internacional de Radio URSI 2000. Zaragoza, España. Septiembre 2000.
- [3] L. Carrera and M. Mora. "License Plate Detection using Neural Networks". 4th International Workshop on Soft Computing Models in Industrial Applications IWANN-SOCO 2009. Salamanca, España. Junio 2009.
- [4] S. Rovetta and R. Zunino. "License-Plate Localization by Using Vector Quantization". 24th International Conference on Acoustics, Speech, and Signal Processing ICASSP 1999. Vol. 2, pp. 1113-1116. Agosto 2002. ISSN: 1520-6149. DOI: 10.1109/ICASSP.1999.759939.
- [5] F. Wang, L. Man, B. Wang, Y. Xiao, W. Pan and X. Lu. "Fuzzy-based algorithm for color recognition of license plates". Pattern Recognition Letters. Vol. 29, pp. 1007-1020. Mayo 2008. ISSN: 0167-8655. DOI: 10.1016/j.patrec.2008.01.026.
- [6] V. Abolghasemi and A. Ahmadyfard. "An edge-based color-aided method for license plate detection". Image and Vision Computing. Vol. 27, pp. 1134-1142. Julio 2009. ISSN: 0262-8856. DOI: 10.1016/j.imavis.2008.10.012.
- [7] J. Jiao, Q. Ye and Q. Huang. "A configurable method for multi-style license plate recognition". Pattern Recognition. Vol. 42, pp. 358-369. Marzo 2009. ISSN: 0031-3203. DOI: 10.1016/j.patcog.2008.08.016.
- [8] J. Rojas y D. Aracena. "Detección y Reconocimiento de patentes vehiculares por Visión Computacional". 13° Congreso Internacional Sudamericano de Sistemas e Informática CISAISI 2009. Arica, Chile. Octubre 2009.
- [9] M. Zahedi and S. Mahdi Salehia. "License Plate Recognition System Based on SIFT Features". Procedia Computer Science. Vol. 3, pp. 998-1002. 2011. ISSN: 1877-0509. DOI: 10.1016/j.procs.2010.12.164.
- [10] D. Lowe. "Distinctive Image Features From Scale Invariant Keypoints". International Journal of Computer Vision. Vol. 60 N° 2, pp. 91-110. Noviembre 2004. ISSN: 0920-5691. DOI: 10.1023/B:VISI.0000029664.99615.94.
- [11] J. B. MacQueen. "Some Methods for classification and Analysis of Multivariate Observations". 5th Berkeley Symposium on Mathematical Statistics and Probability. Vol. 1, pp. 281-297. 1966.
- [12] J. Dunn. "A fuzzy relative of the ISODATA process and its use in detecting compact, well-separated clusters". Journal of Cybernetics. Vol. 3 N° 3, pp. 32-57. 1973.
- [13] K. Fukunaga and L.D. Hostetler. "The estimation of the gradient of a density function, with applications in pattern recognition". IEEE Trans. Information Theory, Vol. 21, pp. 32-40, 1975.