

Reconocimiento automático de objetos y caracteres en una baraja francesa

Walter Sebastián Masi¹, Santiago Germán Silvestre¹, Alfredo Nicolás Campos^{1,2}

¹Universidad Tecnológica Nacional, Facultad Regional Buenos Aires, Departamento de Electrónica, Medrano 951 (C1179AAQ) Ciudad Autónoma de Buenos Aires, Argentina

²Instituto Nacional de Tecnología Agropecuaria, Instituto de Clima y Agua, Av. Rivadavia 1439 (C1033AAE), Ciudad Autónoma de Buenos Aires, Argentina

sebastianmasi@hotmail.com

Recibido el 20 de Septiembre de 2013 y aprobado el 23 de Octubre de 2013

Resumen

El reconocimiento de objetos y caracteres es una tarea básica en la actividad humana. El procesamiento digital de imágenes ofrece la oportunidad de realizar esta tarea de manera automática, objetiva y veloz. Actualmente existen varias metodologías para el reconocimiento de objetos y caracteres, sin embargo ninguna presenta una respuesta que abarque todos los casos posibles. En este trabajo se decidió desarrollar e implementar algoritmos de transformada de Hough, simetría geométrica, correlación espacial contra patrones, detección de color y conteo de elementos, para la determinación del valor y el palo de naipes de baraja francesa. Se aplicaron distintos tipos y niveles de distorsión en las barajas (desenfoque, ruido gaussiano y rotación). Como resultado de aplicar conjuntamente los algoritmos de reconocimiento se obtuvo una efectividad del 96% para naipes en buen estado. En los naipes distorsionados se llegó a valores mayores del 75% siendo los naipes con rotación (principalmente las figuras) los que presentaron mayor error.

PALABRAS CLAVE: RECONOCIMIENTO DE CARACTERES – OCR – RECONOCIMIENTO DE OBJETOS – PROCESAMIENTO DIGITAL – RECONOCIMIENTO DE NAIPES

Abstract

Objects and characters recognition is a common activity on our life. Digital images processing offers high speed automation, objectivity and accuracy. Nowadays there are many objects and characters recognition methodologies. However, none of them shows a solution which could apply for all possible cases. In this paper, pattern comparison, geometric symmetry, Hough transform, color detection and objects count procedures were developed to get the value and suit of cards belonging to a standard French 52-card deck.

Different kind and levels of distortions were used, like unfocused, Gaussian noise and rotation distortions. 96% efficiency was obtained as a result of applying all the recognition algorithms in a deck in good condition. In the case of distorted cards, were reached values greater than 75%. The rotated cards (figures mainly) showed the biggest error level.

KEYWORDS: CHARACTERS RECOGNITION – OCR – OBJECTS RECOGNITION – DIGITAL PROCESSING – CARDS RECOGNITION

Introducción

La visión artificial también conocida como visión por computador, es un subcampo de la inteligencia artificial donde el propósito es lograr que un computador "entienda" las características de una escena. La importancia de este campo de investigación se basa en la posibilidad de automatizar procesos de detección y reconocimiento de objetos en imágenes, los cuales facilitan la evaluación de resultados a gran escala con una capacidad de procesamiento superior a la evaluación por parte de un ser humano. La necesidad de reconocer de forma automática caracteres, formas y objetos tiene un amplio espectro de utilización, en reconocimiento de etiquetas, códigos de barra, presencia de objetos, huellas digitales (Donald, H. y Mc Mahon, 1974, p. 1), escritura manuscrita, entre otros.

La principal problemática que presenta el reconocimiento de objetos es la calidad de la imagen a procesar y las condiciones en las cuales fue adquirida la misma (iluminación no uniforme, sombras, oclusión parcial) (T. Shivanand, 2010, p. 1), (Shan Du, 2013, p. 1), este es un factor importante y determinante para obtener un resultado satisfactorio en nuestro sistema de reconocimiento (Shan Du, 2013, p. 2). Para el Reconocimiento Óptico de Caracteres u OCR, se emplean metodologías como redes neuronales, k vecinos más próximos y contrastación contra patrones entre otros (T. Arrighi, 2012, p. 1). La tecnología OCR proporciona la habilidad de convertir imágenes de caracteres, en caracteres capaces de ser interpretados por una computadora. Así, las imágenes de caracteres en letra de máquina son extraídas de un mapa de bits de la imagen obtenida por un escáner (D. Barragán, s.f., p. 1). Existe la posibilidad de tener ambigüedades en el reconocimiento de caracteres (1/I, 0/O, 0/D, 2/Z, 8/B, and 5/S) utilizando la tecnología OCR (Christos-Nikolaos E. Anagnostopoulos, 2008, p. 11) por lo que se empleó este método cuidadosamente para un reducido conjunto de caracteres aislando las posibilidades de falla.

El objetivo del presente trabajo consiste en la comparación de procedimientos en la detección e identificación de los naipes de sobre diferentes barajas. Para ello se parte de la imagen escaneada del naipe donde se encuentra el superíndice con el valor y el centro de la mis-

ma donde están los elementos o el dibujo de la figura si corresponde, luego procesarlos con distintos algoritmos y arrojar los resultados de las operaciones junto con sus respectivos tiempos de procesamiento, comparando así su performance y efectividad.

Materiales y métodos

Datos utilizados

Una baraja francesa consta de 52 naipes que se agrupan en 4 palos, dos rojos y dos negros. Consta de 13 valores: 9 numéricos (del 2 al 10) y 4 figuras (J, Q, K, A) (Fig. 1).

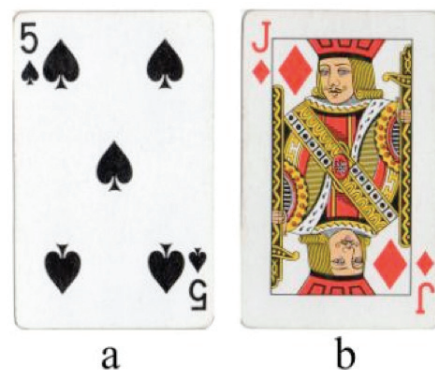


Fig. 1. a) Naipe numérico; b) Naipe figura

Se digitalizaron con una resolución de 300 dpi, dos barajas, distintas entre sí, una de ellas con pronunciado desgaste debido a su intensivo uso.

Metodología

Se le realizaron alteraciones a la baraja en buen estado por medio de filtros digitales para simular condiciones desfavorables. Se utilizó esta diversidad de condiciones para poder contemplar los distintos tipos de distribuciones de formas, colores y caracteres junto con las consecuencias de distintas intensidades de uso y así poder realizar distintas comparaciones de rendimiento y efectividad sobre los métodos propuestos con el fin de evaluar los tiempos de operación junto con su resultado arrojado al momento de identificar el naipe en cuestión. Partiendo de una baraja digitalizada sin alteraciones, la misma fue afectada por los siguientes filtros digitales, ver Figura 2.

. Ruido Gaussiano Monocromático hasta en un 30%. Es un ruido cuya densidad de probabilidad responde a una distribución normal.

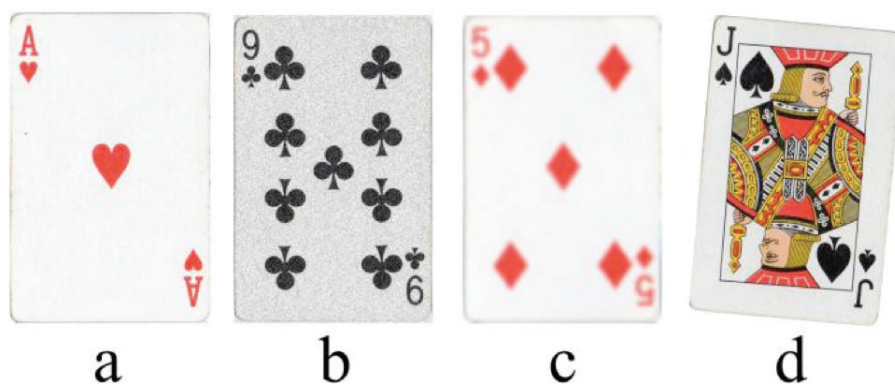


Fig. 2. Alteraciones en los naipes a) Imagen original; b) 30% de ruido Gaussiano; c) Desenfoque Gaussiano de 7 píxeles; d) Rotación horaria de 5 grados

. Desenfoque Gaussiano hasta un radio de 7 píxeles. Este desenfoque aplicado a un mapa de bits mezcla ligeramente los colores de los píxeles vecinos el uno al otro, haciendo que la imagen pierda detalles, y de esta forma se vea más suave y menos nítida.

. Rotación en sentido horario de hasta 5 grados.

La metodología que se utilizó en este procesamiento se puede resumir en el esquema presentado en la Figura 3. (Yao Yuan, 2010, p. 1)

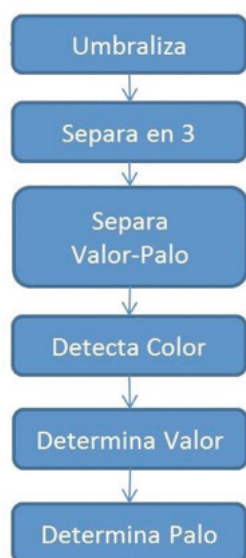


Fig. 3. Metodología utilizada para el análisis

En primera instancia se umbralizó la imagen para la eliminación de la mayor parte del ruido. Luego se la separó en 3 partes, para poder segmentar el valor y el palo. (Fig. 4)

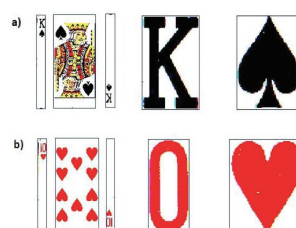


Fig. 4. Naipe separado en 3 partes, luego segmentado el valor y el palo. a) Figura; b) Numérico

Luego que se individualizó el valor y el palo, se detectó el color y si es o no figura. Con esa información se le aplicaron los distintos métodos a los sectores importantes del naipe (centro, valor y palo).

Determinación del Valor

Utilizando la imagen que contiene la información del valor del naipe, se utilizaron diversos métodos para determinación del mismo. Dependiendo de si es o no una figura se utilizaron distintas metodologías para poder tener siempre dos caminos independientes para la obtención del valor (Fig. 5).

Es una figura (J, Q, K, A):

Se analizó la simetría geométrica del segmento que contiene el valor, ya que es distinta para cada una de las cuatro letras. El naipe *Jack* no tiene simetría vertical (Fig. 6.1) en cambio el naipe *Queen* sí (Fig. 6.2). En el caso de la simetría vertical se dividió verticalmente por la mitad la imagen. La parte derecha se reflejó

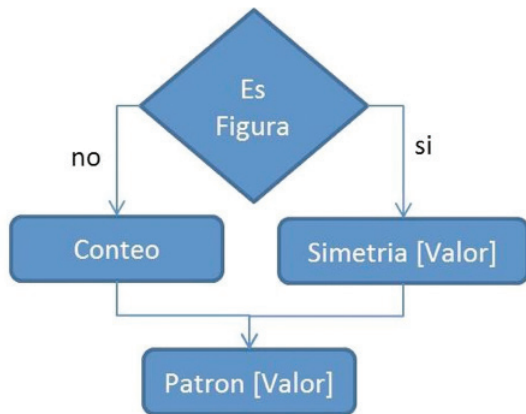


Fig. 5. Análisis del valor

horizontalmente y luego se comparó con la mitad izquierda. En la simetría horizontal se realizó la misma operación pero cortando la imagen horizontalmente.

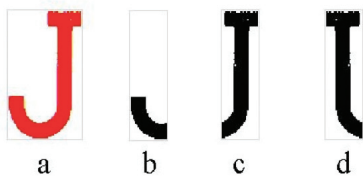


Fig. 6.1. Jack. a) Original b) Mitad izquierda c) Mitad derecha d) Mitad derecha reflejada

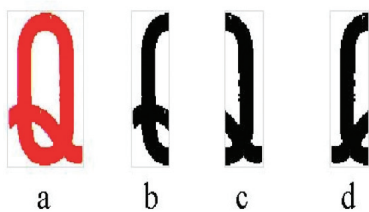


Fig. 6.2. Queen. a) Original b) Mitad izquierda c) Mitad derecha d) Mitad derecha reflejada

Es un número (del 2 al 10):

Para los casos en que era un número (2 al 10) se contaron la cantidad de objetos que se encuentran en el sector medio del naipe. Por ejemplo para el naipe 10 de pica, se contaron 10 elementos (Fig. 7).

Ambos casos (numérico o figura):

Se contrastó contra un patrón el segmento de la imagen que contiene el valor del naipe. Esta comparación se realizó mediante una correlación en dos dimensiones entre la imagen a analizar y

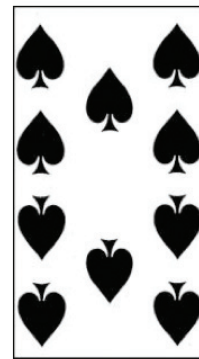


Fig. 7. Sector medio del naipe 10 de pica

un patrón fijo. Esto permite comparar ambas imágenes y obtener una probabilidad de similitud entre ambas.

Determinación del palo

Para determinar el palo se pudieron utilizar entre 2 o 3 métodos según el color fuera rojo o negro (Fig. 8).

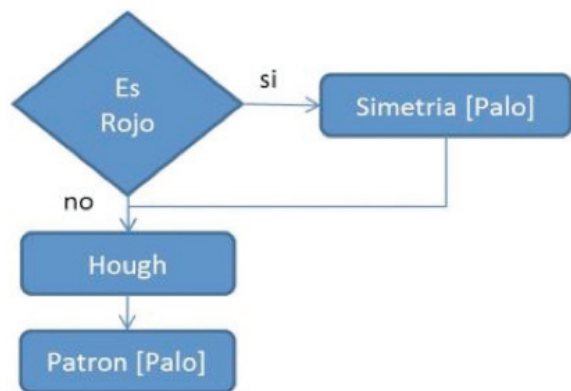


Fig. 8. Análisis del palo

Si el color del naipe es rojo, se analizó la simetría geométrica del palo, ya que el diamante tiene simetría vertical y horizontal y el corazón no.

Independientemente de cuál fuera el color del naipe se aplicó la Transformada Circular de Hough (Ec.1) para identificar las distintas formas características de cada palo (líneas, círculos, etc.).

Esta transformada consiste en recorrer la imagen con círculos de un radio fijo (Fig. 9). Para el centro de cada círculo se suman cuántos puntos del círculo coinciden con la imagen original, de esta manera se obtiene un dominio transformado. Luego el centro con mayor cantidad de coincidencias es el lugar de la imagen donde es más probable que haya un círculo (Fig.10).

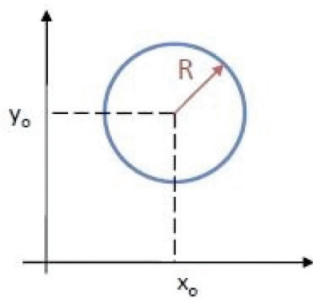


Fig. 9. Representación de un círculo mediante su ecuación 1

Por otra parte, al igual que para determinar el valor, se realizó una comparación contra patrón a través de una correlación en dos dimensiones con un conjunto de imágenes definidas y estandarizadas en tamaño.

Resultados y Discusión

Primera prueba (Baraja N°1):

Se realizó sobre una baraja en muy buen estado. Como resultado se obtuvo un 98% de efectividad. Todos los métodos empleados funcionaron correctamente en todos los naipes analizados. El único naipe en el que un método falló fue el Ace de Picas en el cual la detección del valor por conteo fue incorrecto. El Ace de Picas de esta baraja contiene muchas formas y textos que perjudican el correcto funcionamiento de la detección de la cantidad de elementos en el segmento medio del naipe (Fig. 11). Como el procesamiento empleado utiliza más de un método, la determinación del valor se pudo realizar correctamente a través de la identificación de su superíndice contrastando contra un patrón.

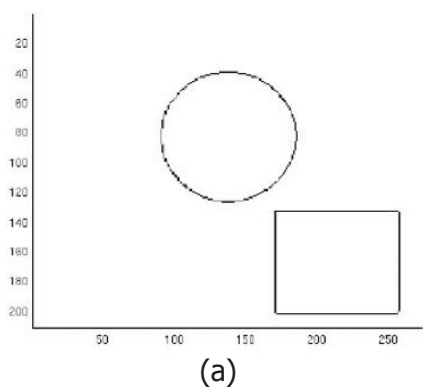


Figura 11. Ace de Picas de la baraja N°1

Segunda prueba (Baraja N°2)

Para esta prueba se utilizó otra baraja completamente distinta. Esta baraja difiere en el tamaño de las figuras, las formas y tamaño tanto de los palos como de los valores respecto de la que se utilizó en la primera prueba. El resultado que se obtuvo fue una efectividad del 83%. La efectividad de cartas analizadas en todos los casos, se tomó como la cantidad de cartas en las que todos los métodos aplicados fueron exitosos.

Al realizar la comparación entre los resultados obtenidos para cada método en cada baraja (Tabla I) se observó que la efectividad en todos los casos fue igual o menor. Esto se debe a que la baraja N°1 fue la que se utilizó como referencia. Esto se aprecia notablemente en los métodos que utilizan comparación contra patrón. En cambio el resto se mantuvieron en niveles cercanos, ya que no dependen de ninguna referencia.

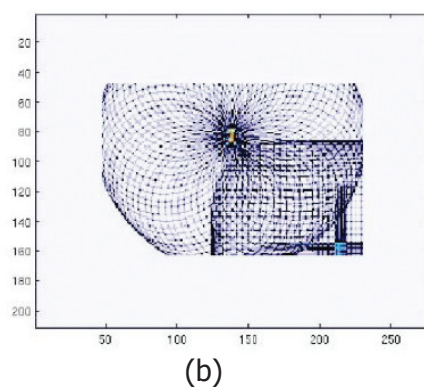


Fig. 10. a) Imagen donde se encuentra un círculo de radio R y otra figura aleatoria; b) Resultado de aplicar la Transformada de Hough

Método	Baraja 1	Baraja 2
Color	100%	100%
Figura	98%	100%
Valor Patrón	100%	89%
Valor Simetría	100%	81%
Valor Conteo	100%	97%
Palo Patrón	100%	92%
Palo Hough	100%	97%
Palo Simetría	100%	100%

Tabla 1. Resultados barajas N°1 y N°2

Tercera prueba con distorsiones (Baraja N°1)

El objetivo fue poner a prueba el funcionamiento de los procesos de reconocimiento en condiciones extremas. Para ello se utilizó la baraja N°1 a la cual se le aplicaron distintos filtrados digitales obteniendo distorsiones que perjudican la determinación de los parámetros.

Se comprobó que los métodos empleados para la separación de las zonas donde se encuentra la información valiosa a analizar funcionaron correctamente para todos los casos extremos de distorsión. (Fig. 12, 13 y 14)



Fig. 12. Resultados del desenfoque Gaussiano



Fig. 13. Resultados de la rotación



Fig. 14. Resultados del ruido Gaussiano

Como se aprecia en las figuras previas, los métodos que dependen de las formas del valor y el palo no se ven afectados por la rotación o el ruido Gaussiano pero sí por el desenfoque, ya que se distorsiona notablemente la forma de la imagen.

En estas pruebas, el desenfoque Gaussiano (*blur*) se aplicó desde un radio igual a 1 píxel hasta un radio igual a 7 píxeles incrementándolo de a un píxel más por prueba. La rotación se aplicó desde 1 grado hasta 5 grados siempre en sentido horario con incrementos de un grado. El ruido Gaussiano monocromático se aplicó desde un 5% hasta un 30% en pasos de a 5%.

Los resultados de efectividad obtenidos para cada distorsión de la baraja N°1 se detallan en la Tabla 2.

Distorsión	Efectividad
blur1	94%
blur2	92%
blur3	94%
blur4	92%
blur5	90%
blur6	77%
blur7	29%
ruido5	96%
ruido10	96%
ruido15	94%
ruido20	98%
ruido25	85%
ruido30	79%
rotacion1	96%
rotacion2	90%
rotacion3	81%
rotacion4	75%
rotacion5	75%

Tabla 2. Resultados de las distintas distorsiones sobre la baraja N° 1

Se pudo apreciar que para 17 de las 18 abarajas los resultados fueron superiores al 75% de cartas en las que ningún método falló.

Relacionando los resultados de eficacia y analizando la Figura 13, se notó que la inclinación empieza a afectar a partir de un nivel de inclinación de 4 grados. Esto afecta notoriamente en la segmentación de la parte izquierda central y

derecha del naipe.

Por otro lado se analizó la efectividad de cada método por separado para cada baraja con distorsión. (Tabla 3)

Distorsión	Color	Es figura	Valor Figura		Valor no Figura		Palo			
			Patrón	Simetría	Patrón	Conteo	Patrón	Hough	Simetría	
bl	blur1	98%	94%	87%	100%	100%	100%	100%	98%	100%
	blur2	98%	100%	80%	100%	100%	100%	100%	98%	100%
	blur3	98%	100%	87%	100%	100%	100%	100%	98%	100%
	blur4	98%	94%	80%	100%	100%	100%	98%	98%	100%
	blur5	98%	94%	73%	100%	100%	100%	98%	98%	100%
	blur6	98%	88%	60%	89%	100%	100%	98%	96%	100%
	blur7	96%	69%	53%	56%	100%	100%	81%	90%	100%
Ruid	ruido5	98%	100%	100%	100%	100%	100%	98%	100%	96%
	ruido10	100%	94%	93%	100%	100%	100%	100%	100%	100%
	ruido15	100%	100%	87%	100%	100%	100%	100%	100%	100%
	ruido20	98%	100%	100%	100%	100%	100%	100%	100%	100%
	ruido25	98%	100%	87%	100%	100%	100%	98%	100%	84%
Ro	ruido30	100%	100%	77%	97%	100%	100%	100%	98%	85%
	rot1	100%	100%	93%	100%	100%	100%	100%	100%	100%
	rot2	98%	88%	80%	100%	100%	97%	100%	98%	100%
Rot	rot3	98%	94%	40%	100%	100%	100%	98%	100%	100%
	rot4	98%	94%	20%	100%	100%	100%	98%	100%	100%
	rot5	96%	94%	20%	100%	100%	100%	96%	100%	100%

Tabla 3. Comparación de resultado para cada baraja con distorsión

En todos los casos analizados la efectividad de los métodos de detección de color y determinación de si el naipe es literal o numeral fue superior al 96%. Esto se debe a que el color rojo y el negro son muy distintos y no dependen de la forma, ni se ven afectados por el ruido. Es importante que estos valores sean altos ya que algunos métodos dependen del color y de si es figura o no.

Los resultados del valor por patrón si es figura fue casi inmune a la rotación y al ruido Gaussiano (efectividad mayor al 88%) ya que no deforman mucho la forma de la letra, en cambio el desenfoque afectó más a este método porque sí lo hace, como se aprecia en la Figura 11.

El ruido Gaussiano en general no afectó a ningún método, en el peor caso la efectividad fue del 77% y en la mayoría fue del 100%. Esta distorsión tuvo poco impacto porque el proceso aplica métodos de eliminación de ruido previo a la aplicación del resto de los métodos.

También se pudo apreciar que la rotación afectó mucho a los métodos de obtención del valor por simetría de las figuras y no a los métodos de simetría de los palos. Esto se debe a que una letra se ve mucho más afectada en su simetría si se la inclina que cualquiera de los palos.

Al mirar los resultados de la determinación del valor por conteo para los naipes que no son figura se notó que es un método altamente efectivo y confiable ya que los ruidos casi no lo afectaron (en todos los casos 100% y en uno 97% de efectividad).

En cuanto a los resultados obtenidos en la determinación del palo mediante el método de la Transformada Circular de Hough es prácticamente inmune al ruido y a la rotación, y es afectada en muy poca medida por un desenfoque Gaussiano de 7 píxeles de radio (siendo este desenfoque un ruido muy importante). En cambio, el método de comparación contra patrón, disminuye su efectividad con altos niveles de desenfoque.

Para el caso de naipes literales, se observó que el método de comparación contra un patrón siempre fue superior al método de simetría geométrica, el cual se ve muy afectado por la rotación y el desenfoque Gaussiano.

Conclusiones

Se pudo concluir que todos los distintos procesos utilizados fueron altamente efectivos, tanto al cambiar de baraja como al agregarle distorsión mediante el filtrado digital a las imágenes originales. Se pueden utilizar los resultados comparativos obtenidos en las pruebas para determinar el valor y el palo del naipe en el caso de que un método no coincida con el otro ya que muestran qué métodos son más efectivos e independientes a los distintos ruidos aplicados, por lo tanto se puede determinar cuál es más confiable.

La obtención del palo fue el proceso con menos fallas, luego la determinación del valor de los naipes que no son figuras y por último el valor de los naipes que son figuras. Por otro lado se observa que el método de patrón fue muy efectivo en todas las pruebas, esto a pesar de que las formas de las letras, números y palos varían considerablemente de una baraja a otra y el patrón siempre es el mismo

A futuro se podrán mejorar los métodos agregando por ejemplo un método que detecte la rotación del naipe, para luego corregirla y así mejorar los resultados obtenidos para la distorsión por inclinación. Analizando mayor cantidad de barajas se podrían mejorar los umbrales utilizados en los distintos métodos (por ejemplo en la segmentación del naipe).

Referencias

- ANAGNOSTOPOULOS, CH.N. E.; ANAGNOSTOPOULOS, I. E.; PSOROULAS, I. D.; LOUMOS, V. y KAYAFAS, E., September 2008, License Plate Recognition From Still Images and Video Sequences: A Survey IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, VOL. 9, NO. 3
- ARRIGHI, T.; ROJAS, J. E.; SOTO, J.C.; MADRIGAL, C. A. y LONDOÑO, J. A., 2012, Recognition and Classification of Numerical Labels Using Digital Image Processing Techniques.
- BARRAGÁN, D. s.f. Reconocimiento de Caracteres Ópticos (OCR) Usando MATLAB.
- CUEVAS JIMÉNEZ, E.V.; ZALDIVAR NAVARRO, D. S.F., Visión por computador utilizando MATLAB y el Toolbox de Procesamiento Digital de Imágenes.
- DU, S.; IBRAHIM, M.; SHEHATA, M. y BADAWY, W., February 2013, Automatic License Plate Recognition (ALPR): A State-of-the-Art Review IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 23, NO. 2.
- JARAMILLO, M.A.; FERNÁNDEZ, J.A. y MARTÍNEZ DE SALAZAR, E. s.f. Filtro de Mediana Sobre Redes Neuronales.
- MCCMAHON, D. H.; JOHNSON, G. L.; STANLEY L. T. y WHITNEY, C. G. , April 1975, A Hybrid Optical Computer Processing Technique, IEEE Transactions on computers.
- SHIVANAND, T.; RAHMAN,S. y PILLAI, G., 2010, Efficient and Robust Detection and Recognition of Objects in Grayscale Images.
- YAO YUAN, WU XIAO-LI. X. I., 2010, Vehicle License Plate Recognition System Based on Digital Image Processing