

Identification of alphanumeric characters in license plates based on image processing with commercial software as a complement or alternative to image processing with Amped FIVE forensic software

Lizbardo Orellano-Benancio¹, Ricardo Muñoz-Canales², Danny Humpire-Molina³, Enrique Lee Huamani⁴

¹ Oficina de Peritajes del Ministerio Público, Perú, orellano01@yahoo.com

² Oficina de Peritajes del Ministerio Público, Perú, ramc2240@gmail.com

³ Oficina de Peritajes del Ministerio Público, Perú, dhumpire@hotmail.com

⁴Image Processing Research Laboratory (INTI-Lab), Universidad de Ciencias y Humanidades, Lima-Perú, ehuanianu@uch.edu.pe

ABSTRACT

The identification of the characters (symbols) that comprise a vehicle license plate number has certain considerations in Peru. Regarding software, there is a series of filters for digital image processing that allow the search for visual evidence in an image or a video. Likewise, at a forensic level, the of Bureau of Forensic Expert Analysis of the Prosecution Authority uses the Amped FIVE forensic software with very good results in the different forensic expert works related to images and videos, which are requested via official letter and the evidence is submitted in a closed and/or sealed envelope with its respective chain of custody. The identification of the characters of the license plate of a vehicle that has been involved in a criminal act is the object of study, with the purpose of finding the owner. Being this research work related to image processing through the use of digital filters with commercial software and its application in the identification of the characters of the license plate of the vehicle as a complement and/or alternative to the Amped FIVE software, using an image of 480 x 320 pixels as evidence that hinders the processing of digital images. Likewise, the details of the methodology and procedure in the work of the expert in digital forensic analysis have been restricted since the research is oriented to the technical procedure of the use of digital image processing and the confidentiality of the forensic expert work.

Key words: Forensic, enhancement, Amped FIVE, image, vehicle license plate, image processing

I. INTRODUCTION

When a criminal act occurs and it is captured in image or video by video-surveillance systems, digital cameras, cell phones, among others, at the request of a prosecutor and for the clarification of the facts, different kinds of forensic expert work are carried out, such as the extraction and securing of the videos, analysis of the operation of the

different equipment, authentication to determine whether an image or video has been subject to editing, manipulation, among others, the enhancement of the image or video and the identification of the license plate, among other kinds of forensic expert work, being the chain of custody: "the procedure to guarantee the individualization, security, and preservation of the material elements and evidences, collected according to their nature or incorporated in all the investigation of a punishable act, destined to guarantee their authenticity, for the effects of the process. The minutes, signs, forms and packaging are part of the chain of custody, which is detailed in Figure 1.

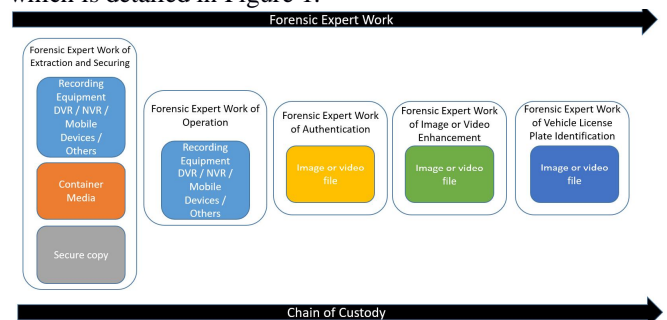


Figure 1: Forensic expert work process

Amped FIVE is a forensic software that allows the processing of images and digital videos, in a single platform, without altering the file under study, being one of the advantages the use of a series of mathematical filters that allow to optimize time in a certain type of forensic expert work. Below, there is an image that will be used to detail part of the advantages of Amped FIVE as shown in Figure 2 and Figure 3.

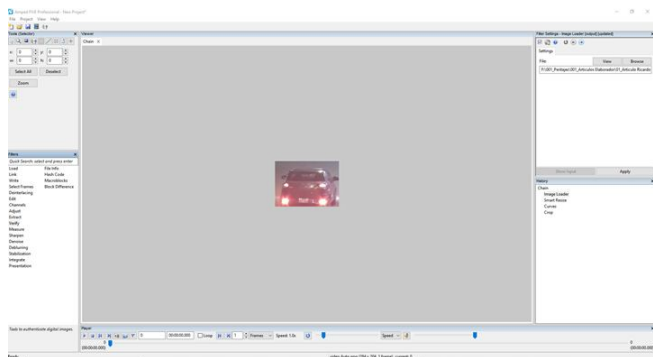


Figure 2: Amped FIVE software

The resize filter, or smart zoom, integrates a series of interpolation algorithms, with one of them "Lanczos" being slightly sharper than the "bicubic" one, but a little less efficient in computational cost, displaying in the following image a resize or smart zoom 6.

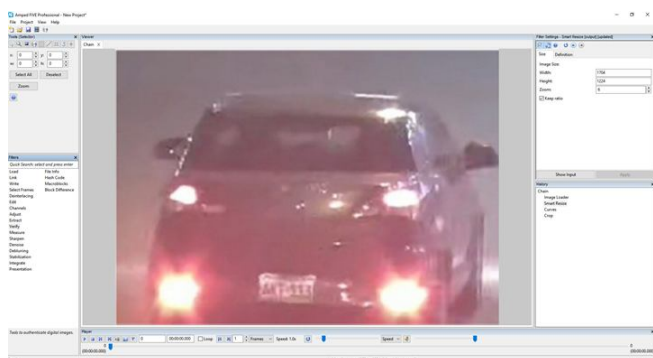


Figure 3: Using the Smart Resize filter (Zoom 6)

The Curves filter, which adjusts the tone values following a smooth curve defined by control points, is used. It supports independent modification of the intensity and RGB channels, as well as multiple interpolation modes detailed in Figure 4 along with the cutting of the area of the vehicle license plate under study.

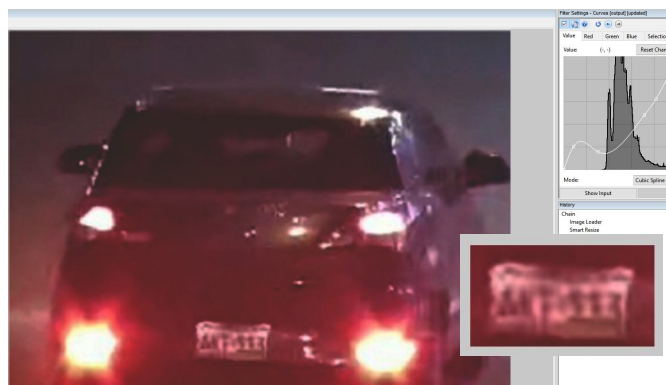


Figure 4: Use of the Curves filter

From the above detailed, the morphological similarities with the alphanumeric characters have been analyzed, concluding that:

- The first character has morphological similarity with the letter A
- The second character has morphological similarity with the letters N and V.
- The sixth character has morphological similarity with the number 3.
- The third, fourth and fifth characters cannot be detailed.

For this research, some digital filters have been used using a commercial software called MATLAB to complement the analysis with Amped FIVE or to determine whether commercial software can optionally be used in the digital forensic analysis of images and videos.

2. PROCEDURE

As part of the procedure for the forensic expert work of vehicle license plate identification and image processing, the current development of the digital filtering process flow is detailed in Figure 5.

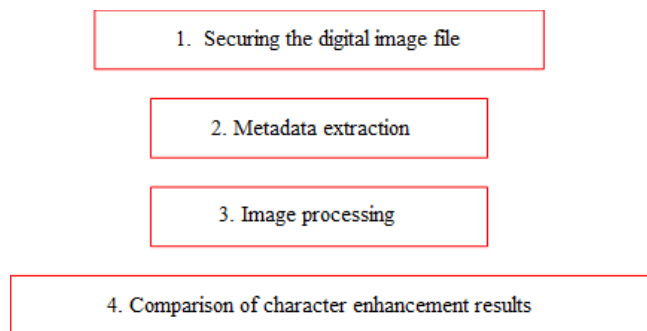


Figure 5: Process flow carried out in this work.

2.1 Securing the digital image file

For this research, we start from securing the image sample in a USB digital storage device (recipient), from which the HASH SHA 256 code [1] of the image file is calculated, with the HashMyFiles v2.23 software, for evidence assurance, as seen in Figure 6.

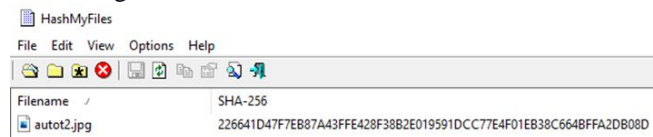


Figure 6: HASH code SHA-256 of the received image file

In order not to alter the original file, we proceed to make a forensic copy and verify its correspondence with the HASH code for digital image processing with commercial software.

2.2 Metadata extraction

At this stage, we obtain the technical information that is recorded in the digital image file. The metadata [3] is shown below:

- File name
- Container format
- Image resolution
- Bit density
- Compression code (codec)
- Color space
- Chroma key subsampling
- File size

We obtain this information to establish the appropriate procedures to be performed for image processing. This information is obtained using the software MediaInfo v19.09 of the company MediaArea.net SARL from the file called "auto2.jpg", as detailed in Figure 7.

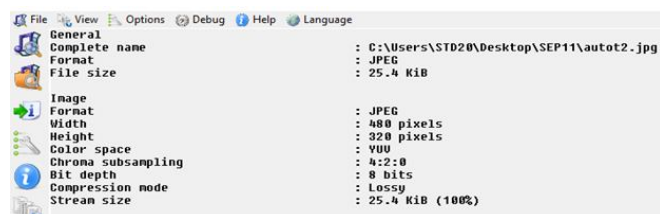


Figure 7: Obtaining metadata from a test audio file with HashMyFiles v19.09 software

2.3 Image processing

At this stage, the goal is to perform mathematical operations in the space domain and in frequency comprised of the numerical values of the pixels in matrix form, this using MATLAB with the forensic copy of the image so as not to alter the original source file.

It is observed in the characteristics obtained from the metadata that the image has a resolution of 480x320 pixels in container format "JPG" and compression mode with loss in bitmap with 8-bit resolution; this shows that the processes of adaptation and improvement of the image will be required by the expert to identify the alphanumeric characters of the vehicle, as detailed in this work.

In the image received, there is a vehicle with the plate of which we seek to identify the alphanumeric characters, as shown in Figure 8.



Figure 8: Image load of the investigated vehicle

The first process to perform will be the function of "cropping" this to delimit the section of the image where the information of interest is located, which is the license plate of the vehicle; in Figure 9, there is a capture of the vehicle license plate.



Figure 9: Capture of the vehicle license plate

Afterwards, we proceed to "rescale" the image using an interpolation algorithm, in this case the "Lanczos" method as one of the techniques described in [2], whose representation is shown in equation (1).

$$u(x, y) = \frac{\sum_{m,n} v_{m,n} K_1(x-m) K_1(y-n)}{\sum_{m,n} K_1(x-m) K_1(y-n)} \quad (1)$$

The scaling process is shown with a factor of 8 resulting in an image of 3840 x 2560 pixels, thus increasing the resolution of the image, softening the edges of the objects that have been affected by the scaling. Figure 10 shows the vehicle license plate image scaled with the Lanczos algorithm.



Figure 10: Image of the license plate scaled in magnitude 8 (3840 x 2560 pixels) with the Lanczos algorithm

Subsequently, the representation of the RGB color image is converted to grayscale, and the frequency response of the high-pass filter is applied using 2-D convolution [5], [6], from equation (2).

$$C(j, k) = \sum_p \sum_q A(p, q)B(j - p + 1, k - q + 1) \quad (2)$$

The 2-D convolution process is performed with one matrix represented by the image of the license plate and the other matrix represented by the high pass filter coefficients; a test was performed with a low pass model, but the high pass showed better detail; the filtering result is shown in Figure 11.

h2 =

0.8889	0.8889	0.8889
0.8889	0.8889	0.8889
0.8889	0.8889	0.8889

Figure 11: (i) Matrix of high-pass filter coefficients

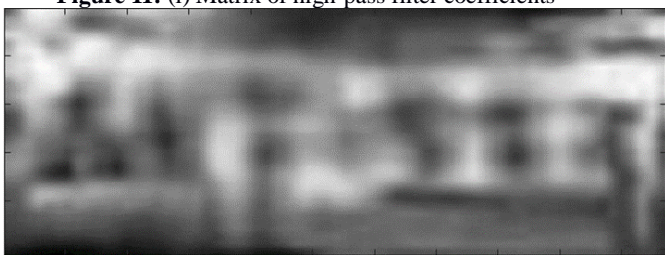


Figure 11: (ii) Filtered image

To improve the details of the image obtained, we proceed to apply a blur masking [7] with a standard deviation of 10, and a blur force of 15, which will refine the edges of the alphanumeric characters to be identified, as shown in Figure 8.



Figure 12: Image of the license plate after applying a blur mask

The pixel values are limited with the binary process with a level of 0.6, this allows to perform the processes of expansion and erosion in the image space, in addition to the inversion of binary image values.

An eroded image is obtained to soften the noise near the edges of the alphanumeric characters in the image with a linear structural vector. A dilation is subsequently performed on the residual image of the alphanumeric characters in the image with a spherical-shaped vector to smooth the edges of the image, as detailed in Figure 13.



Figure 13: (i) Above, eroded binary image.



Figure 13: (ii) Below, dilated binary image.

Finally, the pixel values are reversed to recover the original binarization of the processed license plate image. The image is segmented according to the possible alphanumeric characters to be identified.

In accordance with Law No. 27181 [4], in the agreements specified therein and its Title II called "Codification System of the Unique National License Plate", the standard license plate number for light and heavy vehicles is detailed as follows:

- a) The first character shall be a letter of the alphabet, which identifies the registration area of the vehicle.
- b) The second and third characters shall be alphanumeric digits.
- c) The last three characters shall be expressed in numeric values.

It is important to emphasize that the above is shown graphically in figure 14.



Figure 14: License plate model for light and heavy vehicles

From Figure 14, a character-by-character comparison of the license plate characters contained in red rectangles is made in the enhanced image, as shown in Figure 15.



Figure 15: Image where possible characters to be identified are highlighted

The identification will be conducted with the first, second and sixth character that have a higher number of pixels that comprise their morphology; a cross-correlation [3] with the characters of congruent morphology is conducted.

The cross-correlation function is implemented from the formula of (3), where (*f*) is the image, (*t*) the mean of the image of the model to be compared and (*f_{u,v}*) is the mean of *f*(*x*, *y*). As described (Lewis).

$$\gamma(u, v) = \frac{\sum_{x,y} [f(x, y) - \bar{f}_{u,v}] [t(x-u, y-v) - \bar{t}]}{\left\{ \sum_{x,y} [f(x, y) - \bar{f}_{u,v}]^2 \sum_{x,y} [t(x-u, y-v) - \bar{t}]^2 \right\}^{0.5}} \quad (3)$$

The cross-correlation procedure is performed as follows: an image of 91 pixels wide by 121 pixels high is taken, and compared with a scale image of the same "A" font size of FE-Schrift, which is the alphanumeric character font used for vehicle license plate design [4], as shown in Figure 16.



Figure 16: FE-Schrift font used on license plates in the Peruvian Vehicle System.

3. RESULTS

For our identification analysis, we will use the cross-correlation function of a matrix (processed image), to quantify the similarity of the characters on the license plate with the alphanumeric characters of the first, third and sixth characters. The correlation value of the processed image of the first character with itself, with a resolution of 91 pixels wide by 127 pixels high, results in a peak of the cross-correlation value of the real value 1.0, and when the first processed character is compared with the character "A" of the vehicle license plate character font with the same resolution, it gives as real value 0.3917; this is due to the noise present in the morphology of the processed image. Finally, the value of the correlation with character "4" is shown, which in morphology is the closest to character "A", obtaining the value 0.3554. The images used are shown in Figure 17.



Figure 17: (i) Morphological display of the "A" symbol from the processed image









Figure 17: (ii) izquierda, Imagen procesada. Derecha, carácter "A".



Figure 17. (iii) Left, processed image. Right, character "4". The calculation of the cross-correlation of a matrix (two dimensions) with the third and sixth character is made, which provides sufficient information to make a comparison with characters of similar morphology, as detailed in Table 1.

Table 1: Example of morphological characters

No.	IMAGE CHARACTER	CHARACTER WITH CONGRUENT MORPHOLOGY	RESULT OF MAXIMUM CROSS-CORRELATION
1		A	0.3917
2		V	0.5147
3		It is not possible to specify.	Does not apply.
4		It is not possible to specify.	Does not apply.
5		It is not possible to specify.	Does not apply.
6		3	0.3634

4. DISCUSSIONS AND CONCLUSIONS

From the observations on the processed image referred to as enhanced, it is possible to identify characters with congruent morphology to compare in order to identify the alphanumeric characters of the license plate of the vehicle investigated and establish comparison techniques to quantify the similarity of

the image sectioning. For each alphanumeric character with a specific resolution (91 x127 pixels) in this work, cross-correlation was used in two dimensions; this allows to establish indicators in forensic investigation in the analysis by digital signal processing to identify a vehicle based on its license plate at the crime scene. It was observed that it is necessary to establish additional techniques to capture each character individually and soften the noise at the edges, as well as to modify the viewing angle of the captured license

plate in the image and, thus, obtaining a higher correlation value to be able to differentiate characters of similar morphology, such as the number "4" and the letter "A", considering the noise resulting from the enhancement.

The image processing for the enhancement of visual characteristics to identify the license plate is carried out in a HASH-secured copy of the digital evidence received, therefore, the original source file is not altered or modified and is processed in the memory with the use of the analysis software, where finally the enhancement result is saved in a separate and secure copy.

Likewise, it is concluded that digital image processing with commercial software can complement the "Forensic Expert Work of Imagen or Video Enhancement" [8], and the processing carried out with Amped FIVE, and it is also found that it can be an alternative to it, which is detailed in the results with commercial software and Amped FIVE.

REFERENCES

1. Computer Security Resource Center. (August 1, 2020). SECURE HASH STANDARD. Retrieved from <https://csrc.nist.gov/csrc/media/publications/fips/180/2/archive/2002-08-01/documents/fips180-2.pdf>
2. Getreuer, P. (2011). Linear methods for image interpolation. *Image Processing On Line*, 1, 238-259.
3. Lewis, JP (2005). Fast normalized cross-correlation, *Industrial Light and Magic*. Unpublished.
4. Law No. 27181. Official Gazette El Peruano, Lima, Peru, October 8, 1999.
5. McAndrew, A. (2004). An introduction to digital image processing with matlab notes for scm2511 image processing. School of Computer Science and Mathematics, Victoria University of Technology, 264 (1), 1-264
6. Oppenheim, AV, Willsky, AS, & Nawab, SH (1998). *Signals and systems*. Pearson Education.
7. Singh, H., & Sodhi, JS (2013). Image enhancement using sharpen filters. *International Journal of Latest Trends in Engineering and Technology (IJLTET)*, 2 (2), 84-9
8. A. Lazarte, L. Orellano, W. Gutiérrez, J. Soto, H Rodríguez "Documento Interno Ministerio Público- Guía Análisis Digital Forense" Page. 14, 2020.