MOPM
MEXICAN OPTICS AND
PHOTONICS MEETING 2012

Proceedings

Proceedings of the Mexican Optics and Photonics Meeting 2012 (MOPM 2012)

© 2012 Academia Mexicana de Óptica

www.amo-ac.mx


Edited by
Amalia Martínez García
Eric Rosas
A rotation invariant correlation using an adaptive methodology

Alfredo Solís Ventura, Josué Álvarez Borrego

Centro de Investigación Científica y de Educación Superior de Ensenada.

asolis@cicese.edu.mx, josue@cicece.mx

ABSTRACT:
Distortions of an object due to displacement or rotation, adds difficulty or makes impossible the identification of the object. A low computational cost digital algorithm was developed to recognize images even when these images appear rotated, displaced and immersed in noise. Vectorial signatures are generated from binary masks of concentric rings when this is multiplied by the modulus of the Fourier transform of the image. In this application, vectorial signatures are compared using nonlinear correlations. It was possible to get the identification of the 26 alphabet letters, where each of these were rotated degree by degree from $0^\circ$ to $359^\circ$ generating 9360 different images. Results with a 100% confidence level in the identification of the 26 alphabet letters are shown.

Key words: Image recognition, nonlinear correlation, rings mask, one-dimensional signatures, Fourier transform.

REFERENCES AND LINKS

1. Introduction.
Recognition of images in real time is not an easy issue because the object to recognize can be displaced and to have distortions from rotation or from different kinds of noise. Therefore it is necessary the development of specific techniques to recognize the same object with invariance to a concrete distortion or a set of these. Since 60’s, the binary masks are used to take frequency samples from the diffraction pattern and to get the invariance to position [1].

The rotation invariance is achieved using a binary mask of concentric rings [2]. Recent investigations have used adaptive rings mask, getting a robust one-dimensional vectorial signature, taken only the most important frequencies from the image diffraction pattern. A k-law nonlinear filter is used to realize the digital invariant correlation that gives us information on the similarity between different images [6].

1. Methodology.
The digital algorithm proposed generates an adaptive binary rings mask from the real part of the Fourier transform (FT) of the image. The mask takes samples of the most important frequencies from the modulus of the Fourier
transform of the image to generate a one-dimensional vectorial signature. Signatures are compared using an adaptive
k-law nonlinear correlation.

2.a. **Binary rings mask of the image.** The construction of the binary rings mask starts taking the positive and real
part of the FT associated to a target image. From this reference, adaptive concentric rings mask are generated. Next,
a point to point product with the rings mask and the modulus of the Fourier transform (|FFT|) of the target image was
done. From this product, concentric rings with sampled values from |FFT| are calculated. The next step is to compute
the sum of the intensity values in each ring and to obtain a one-dimensional vectorial signature. This signature
represents a target image with invariance to rotation and position. Same procedure is made for the problem image
(image to be recognized or classified).

2.b. **Nonlinear correlation.** In order to recognize the target, the signature of the problem image (PI) is compared
with the signature of the target image (TI) using the nonlinear correlation, $C_{NL1}$,

$$
C_{NL1}(PI, TI) = PI \otimes TI = \text{FFT}^{-1}\left(\left|\text{FFT}(PI)\right|^k e^{i \delta_{PI}} \left|\text{FFT}(TI)\right|^k e^{-i \delta_{TI}}\right),
$$

where $\otimes$ means correlation, $i = \sqrt{-1}$, $\delta_{PI}$ and $\delta_{TI}$ are the phase of the Fourier transform of the problem image and
the target image signature, respectively, $0 < k < 1$ is the nonlinear coefficient factor [10].

The area value of an image can be used as invariance to rotation; this invariance is obtained because even if the
image is rotated this value will be conserved [9]. In this case, the mean value of the images as parameter invariant to
rotation was used. In order to increase the discrimination between the target and the problem image, a ratio of the
mean values of the images is introduced to Eq.(1).

$$
\text{IF } \text{mean}(PI) < \text{mean}(TI) \text{, ratio} = \frac{\text{mean}(PI)}{\text{mean}(TI)} \text{ ELSE ratio} = \frac{\text{mean}(TI)}{\text{mean}(PI)},
$$

Then the Eq.(2) ratio is introduced to Eq.(1), and has a new adaptive nonlinear correlation, $C_{NL2}$,

$$
C_{NL2}(PI, TI) = PI \otimes TI = \text{FFT}^{-1}\left(\left|\text{FFT}(PI)\right|^{\text{ratio} k} e^{i \delta_{PI}} \left|\text{FFT}(TI)\right|^{k} e^{-i \delta_{TI}}\right),
$$

2. **Results.**
An optimal value of $k$ (ratio-k) to use in Eq.(3) in the presence of additive Gaussian noise and salt and pepper noise
was found. The 26 alphabet letters were analysed, each of these images in Arial font type was rotated degree by
degree from $0^\circ$ to $359^\circ$ generating 9360 different images in black and white format, size of 128 x 128 pixels. 360
different correlations to each letter were computed. The mean value of these correlations are represented in a box
plot, at two standard errors ($\pm 2SE$), the confidence level was 95.4%. If the correlation of different letters mean
correlation values don’t overlap, then we get a confidence level of 100%. In all the cases the identification of the
target images was a success with a confidence level of 100%. The computer time per image is about 0.2 sec.

4. **Conclusion.** The proposed adaptive methodology is able to extract the necessary information from an image to get
a one-dimensional vectorial signature invariant to position and rotation; and using the aid of a novel proposed variant
of a k-law nonlinear correlation to achieve an adaptive version of nonlinear correlation, this methodology is able to
represent a target image with success and excellent results even when the problem image is immerse in noise. It was
possible to get the identification of the 26 alphabet letters, with a confidence level of 100% and using a minimal time
of computational cost in all the cases.

Acknowledgements
This document is based on work partially supported by CONACyT under Grant No. 102007.