Adaptive nonlinear correlation with a binary mask invariant to rotation and scale

Alfredo Solís Ventura a, Josué Álvarez Borrego a,*, Selene Solorza b

a CICESE, División de Física Aplicada, Departamento de Óptica, Carretera Ensenada-Tijuana, No. 3918, Fraccionamiento Zona Playitas, Ensenada, B.C. C.P. 22860, Mexico
b UABC, Facultad de Ciencias, Km. 103, Carretera Tijuana-Ensenada, Ensenada, B.C. C.P. 22860, Mexico

A R T I C L E   I N F O

Article history:
Received 18 September 2014
Received in revised form
12 November 2014
Accepted 13 November 2014
Available online 17 November 2014

Keywords:
Fourier optics
Scale transform
Nonlinear correlation
Binary mask
Pattern recognition
Computer vision

A B S T R A C T

A new adaptive methodology for pattern recognition invariant to scale, rotation and translation is presented. In order to get rotation invariance, this methodology generates an adaptive binary rings mask taking information from the 2D separable scale transform of the target image. With this mask a vectorial signature is generated getting samples from the scale transformation of the target image and it does the same for every problem image. Finally, by the use of a new adaptive nonlinear correlation, the comparison of the signatures of the target and the problem image were done. The performance of this methodology has been experimental verified with a statistical analysis, to know the mean correlation confidence level in the classification of 30 different phytoplankton images in gray scale with rotations from 0° to 180° and ±10% of maximum scale variation with respect to the target image.

1. Introduction

Distortions due to scale changes and orientation of an object make difficult its identification, because these distortions can generate an infinite number of different images of the same object. In order to overcome part of these problems, rotation invariance using a binary mask of concentric rings is used. In order to take only the most important frequencies from the image diffraction pattern, in recent investigations have been used adaptive rings mask [1,2], getting a robust vectorial signature methodology invariant to rotation. These adaptive rings masks were constructed using the real or imaginary part of the Fourier transform of the target image. Several alternative ways to generate the binary rings masks have been studied and all of them show rotation invariance [2].

In this paper a methodology using binary rings masks via the separable implementation of 2D scale transform is obtained. The discrimination between images is done by a new adaptive \( k \)-law nonlinear correlation. In this adaptive nonlinear correlation, the product between the \( k \) constant and the ratio of two invariant indexes is used. These two indexes are obtained from the area and volume of the target image and the problem image. With the use of the adaptive nonlinear correlation, the system increases its capacity to discriminate the pattern in the images.

The material of this work is organized as follows: in Section 2, the invariant adaptive methodology to position, rotation and scale is described, in Section 3, computer simulations are presented and in Section 4, the conclusions are given.

2. Invariant adaptive methodology to position, rotation and scale

The algorithm proposed in this work generates an adaptive binary rings mask from the modulus of the separable 2D scale transform of the target image. In addition, this binary mask takes samples from the modulus of the 2D scale transform of each problem image to generate its vectorial signature.

2.1. Scale transform

Scale is a physical attribute of a signal just like frequency. For a given signal, the frequency content via the Fourier transform can be determined; analogously, a scale transform is needed to indicate the scale content in the signal [3]. Applications of scale concept have been made in speech analysis [4], processing of biological signals [5], machine vibration analysis [6] and other areas [7,8]. The scale transform was also applied in
The scale transform can be written like
\[
D(c) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) \exp(-jc \ln x) \frac{dx}{\sqrt{x}},
\]
(1)

which shows that it is the Mellin transform [11] with the complex argument \(jc + 1/2\). A practical realization of the Mellin transform is given by a logarithmic mapping of the input scene that is, defining the signal by
\[
f_l(x) = \frac{1}{\sqrt{x}} f(\ln x),
\]
(4)

then, substituting (4) into (1), we have
\[
D_l(c) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) \exp(-jcx)dx.
\]
(5)

Let \(F\) be the symbol for Fourier transform, then \(F(c) = D_l(c)\), hence exists a direct relationship of the Mellin and the Fourier transform.

From Eqs. (4) and (5) one can see the scale by resampling the samples uniformly distributed of \(x\) with a logarithmic function. The simplest form to construct multidimensional scale transforms is through the 1D scale transform case. In the 2D case, we can define two feasible constructions, namely, separable and no separable implementations. The former corresponds to the successive application of 1D transform in the XY Cartesian coordinates while the latter is based in a warping operation, converting the original spatial image onto a Cartesian polar-logarithmic coordinates. The direct and inverse 2D scale transform is given by

---

**Fig. 1.** Procedure to obtain the binary rings mask.

**Fig. 2.** Scale transform digitalization: (a) An image of 256 × 256 pixels with a “T Arial Letter”. (b) Modulus of the separable 2D scale transform corresponding to “T Arial Letter”. (c) Profile \(z(c_y) = |D_2(129, c_y)|\). (d) Slope digitalization from \(z(c_y)\), where \(c_y = 1, \ldots, 128\).

**Fig. 3.** Binary ring mask for the “T Arial Letter”.

---
\[ D(c_x, c_y) = \frac{1}{\sqrt{2\pi}} \int_0^\infty \int_0^\infty f(x, y) \frac{\exp \left( -j c_x \ln x - j c_y \ln y \right)}{\sqrt{\lambda y}} \, dx \, dy. \]

\[ f(x, y) = \frac{1}{\sqrt{2\pi}} \int_0^\infty \int_0^\infty D(c_x, c_y) \frac{\exp \left( j c_x \ln x + j c_y \ln y \right)}{\sqrt{\lambda y}} \, dc_x \, dc_y. \]
Let 
\[ x = \exp(x'), y = \exp(y'), \]
then, substituting in Eq. (6),
\[ D(c_x, c_y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\exp(x'), \exp(y')) \exp\left(-jc_x x' - jc_y y'\right) \exp\left(\frac{x' + y'}{2}\right) dx' dy'. \]

2.2. Binary rings mask

The binary mask is used to obtain a sample of the image diffraction pattern and get the advantage of compressing the data of the images. A binary mask can be a set of radial lines in order to achieve the invariance to scale [12] or a set of concentric circles in order to have the invariance to rotation [12,13] or a combination of both like the ring-wedge detector that have invariance to scale and rotation [14]. Recently, a digital nonlinear correlation methodology that uses the concept of an adaptive concentric rings mask to sample the most significant frequencies from a diffraction pattern shows a significant improvement in invariance to rotation [1].

In step 3 the one variable binary function \( z(c_y) \) is calculated like
\[ z(c_y) = \begin{cases} 1, & m < 0 \\ 0, & m \geq 0 \end{cases}, \quad c_y = 1, ..., C - 1, \]

next in step 4, taking \( c_y = C \) as the rotation axis, the graph of \( z \) is rotated 360°s to obtain concentric cylinders of height one, different widths and centered in \((C, C)\) pixel. Finally, taking a cross-section of those cylinders the binary rings mask associated to image TI is built.

In Fig. 2a the target image “T Arial Letter” is a binary image of 256 pixels, where \( C = 129, 1 < x < 256, 1 < y < 256 \) and, \( c_x \) and \( c_y \) represents the transformed axis, \( c_y = 1, ..., 256, ... \)

Fig. 6. Procedure of the invariant image recognition through the 2D separable scale transform and an adaptive correlation method.
2.3. The signature of the image

The procedure to obtain the signature of the image, which is invariant to position, rotation and scale, is described in this section. The first step is set the image, for example the Arial Letters in Fig. 4. Then the modulus of the scale transform of the image is calculated (step 2 in Fig. 4). So, the mask (step 3) is applied in the scale domain to sampling the scale pattern of the object (step 4) by a point to point product of the concentric rings mask and the modulus of the 2D scale transform of the target image. The final step is calculate the sum of the values in each ring to obtain a vectorial signature (step 5 in Fig. 4). This signature represents the image and it is invariant to position, rotation and scale.

Some images could present some distortions due to rotation or scale transformations, for example, in Fig. 5 are showed in solid line the vectorial signatures for the "T Arial Letter" image without distortion and in dotted line the signatures for the image with distortions. Fig. 5a shows in dotted line different vectorial signatures for the "T Arial Letter" image with distortions due the rotation, there is a variability in these signatures with respect to the signature without distortions because the saw tooth effect from the rotated image is incorporated in the signature of the rotated images and the effect due to scale changes are presented in Fig. 5b with a minimal variability, an analogous exercise was done for the "O Arial Letter" image (Fig. 5c and d).

2.4. Adaptive nonlinear correlation

Nonlinear filter is used in this work due to its superior performance in comparison with linear filter techniques in terms of discrimination capability, correlation peak sharpness, and noise robustness [15]. In order to recognize the target into a problem image, the signature of the problem image (SPI) is compared with the signature of the target image (STI) using the k-law nonlinear correlation [16],

\[
\phi_{SPI} = \frac{1}{1 + |F(S_{PI})|^{k} \exp(i\phi_{STI})|F(S_{PT})|^{k} \exp(-i\phi_{STI})}
\]

(13)

where, \(\phi_{SPI}\) and \(\phi_{STI}\) are the phase of the Fourier transform for the signatures of the problem image and the target image respectively, \(0 < k < 1\) is the nonlinear coefficient factor and controls the strength of the nonlinearity [15]. In this work, the \(k\) value is \(0.1\).

A modification to (13) is proposed to generate a new adaptive nonlinear correlation equation, which is more robust in its discrimination performance. Here,

\[
\text{Index}_1 = \frac{\text{volume}(TI)}{\text{area}(TI)},
\]

(14)

\[
\text{Index}_2 = \frac{\text{volume}(PI)}{\text{area}(PI)}.
\]

(15)
are used into the new variable called \( \text{ratio} \),

\[
\text{ratio} = \begin{cases} 
\text{Index}_2, & \text{if } \text{Index}_2 < \text{Index}_1, \\
\text{Index}_1, & \text{if } \text{Index}_1 < \text{Index}_2.
\end{cases}
\]  

(16)

which modifies the form to scale the modulus of the Fourier transform of the problem image signatures in the nonlinear

\[
C_{NL1} (13), \text{ that is }
C_{NL2} = F^{-1} \left[ |F(S_{TI})|^{\text{ratio}} \exp \left( i\varphi_{SN} \right) |F(S_{PI})|^{\text{ratio}} \exp \left( -i\varphi_{SR} \right) \right]
\]  

(17)

2.5. The method

Fig. 6 shows the procedure of the pattern recognition system invariant to position, rotation and scale, through the 2D separable scale transform and an adaptive correlation method. In the first and second step the target image and the problem image are centered with respect to their center of mass. The binary ring mask is generated from the target image, step 3. Next, the modulus of the scale transform for the TI (steps 4) and the PI (step 5) are obtained, called \( ID_{TI} \) and \( ID_{PI} \), respectively. Those modulus images are filtered by the binary rings mask (steps 6 and 7). The \( * \) means a point to point multiplication of the images with the binary rings mask. In step 8, the sum of the sampled values located in each ring represent a ring index value of the binary rings mask to build the vectorial signature of the target image (step 10). The same procedure is performed to the selected problem image (step 9); this image could be any image different or similar to the target, in that manner, the signature (step 11) could be similar or different of the signature of the target.

In step 12, \( \text{index}_1 \) and \( \text{index}_2 \) are calculated. These are the ratio of the area and the volume of the target and the problem images, respectively. The variable ratio takes values less than the unit and it is used to complete the adaptive version of the \( k \)-law correlation Eq. 17. The adaptive nonlinear correlation between the target and the problem image is done in step 14.

3. Computer simulations

To evaluate the performance of the digital system, 30 different species of phytoplankton (real images) were used (Fig. 7). The gray scale images are 320 × 320 pixels. The images were rotated 180° in increments of 10° and scaled from 90% to 110% in increments of 1%, hence a bank of 11,970 problem images was obtained.

Phytoplankton are of great ecological significance, since they constitute the greatest portion of primary producers in the sea.
Microscopy is the principal method used to identify and count phytoplankton. Microscopic counts have been used to describe phytoplankton communities and their spatial and temporal distribution patterns, and have also been used to convert phytoplankton numbers to biomass or energy. However, microscopic examination entails considerable time and labor for the processing of a large number of samples; moreover, mistakes are easily made. Methods enabling the rapid identification and quantification of organisms in a phytoplankton sample are needed.

Solorza and Álvarez-Borrego [1] showed that the linear filters do not work well in the discrimination between objects, getting confidence levels of 68.3% in some cases, because they are not distortion tolerant, but when nonlinear filters are used in the system, the results are excellent because they are less sensible to distortion, increasing the confidence level of the system at least to 95.4% in all cases. Table 1 presents the efficiency of the nonlinear correlation (13) and (17), it is shown that the confidence level of the adaptive nonlinear correlation (17) always is better or equal than (13).

The signatures of the problem images are compared with the signature of the target image using a nonlinear correlation (Fig. 6), if the maximum value of the magnitude for the correlation is significant, hence, the problem image contains the target, otherwise has an image different to the target. The results were boxplotting by the mean correlation with two standard errors (2SE). Fig. 8 shows the results using IM01 like target image and CNL1 (13) like correlation method. Overlap between the mean of the maximum values of correlation results of IM01 and IM05 is presented in this boxplot. Fig. 9 shows the results using the same image as target, but now implementing the adaptive nonlinear correlation CNL2 (17), in this case a very good separation among the correlation values of the target image and the different problem images is
shown. Fig. 10 shows results when IM03 is used as target image and $CNL_1$ (13) as correlation method, in this case a 100% of confidence level is obtained. Fig. 11 takes IM03 as target image but now implementing the adaptive nonlinear correlation $CNL_2$ (17), showing a 100% of confidence level but with a better performance (compared with Fig. 10) because the mean of the maximum values of the problem images correlation goes far away from the mean of the maximum values of the target correlation results.

4. Conclusion

In this work, a method to construct binary mask of concentric rings using the slopes of a profile from the modulus of the separable 2D scale transform was presented. Of this way the binary mask is able to takes the most important data from the modulus of the scale transform. Therefore, it was shown that utilizing the scale transform of the images, the proposed method generates a vectorial signature invariant to position, rotation and scale.

Experimental results of the identification of 30 different species of phytoplankton with variations in the scale and rotation, show that the nonlinear correlation method ($k$-law) achieves high recognition rate between vectorial signatures, with a level of confidence at least of 95.4%. The results also show that the proposed adaptive nonlinear correlation (17) is more robust in its discrimination performance, increasing the level of confidence to 100% in most of the 30 different images to identify.

Acknowledgment

This document is based on work partially supported by CONACyT under Grant nos. 169174, SEP-2008-103898 and CB-2011-01-167361. Alfredo Solís-Ventura is a student in the Ph.D. program of Optics Department in CICESE and supported by CONACyT’s scholarship.

References


Alfredo Solís-Ventura is a Ph.D. student in CICESE in the Applied Physics Division, Optics Department. He received his BS and MS degrees in Electronic Engineering and Optics in 2001 from UABC and 2012 from CICESE, respectively.
Josué Álvarez-Borrego received his BS degree in physical oceanography in 1981 from the Facultad de Ciencias Marinas, Ensenada, Baja California, México, and his MSc and PhD degrees in optics in 1983 and 1993 from CICESE, México. He is a Professor with the Applied Physics Division, Optics Department, CICESE, México. His research interests are image processing with applications to biogenic particles and image processing of the sea surface. He is the author of several scientific papers and presentations at scientific conferences. He is a member of the Mexican Academy of Optics, National Research System (SNI), and the Mexican Sciences Academy.

Selene Solorza received her M.Sc. degree in Earth Sciences from Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE), México, in 1999, and her Ph.D. degree in Earth Sciences from CICESE in 2005. She is a researcher of the Math Department in the Faculty of Sciences, Universidad Autónoma de Baja California (UABC). Her research interest includes computational filters, wave propagation in poroelastic media and, finite difference and finite element schemes applied to geophysics and physics.