Nonlinear filter for pattern recognition using the scale transform

Angel Coronel-Beltrána and Josué Alvarez-Borrego

aDepartamento de Investigación en Física, Universidad de Sonora, Calle Rosales y Blvd. Luis Encinas S/N, Apdo. Postal 5-088, Hermosillo, Sonora, C. P. 83000, México; aUniversidad Autónoma de Baja California, Facultad de Ingeniería, km 103 Carretera Tijuana-Ensenada, Ensenada, B. C., México, C. P. 22860; bCentro de Investigación Científica y de Educación Superior de Ensenada, División de Física Aplicada, Dept. de Óptica, km 107 carretera Tijuana-Ensenada, Ensenada, B.C., México, C. P. 22860

ABSTRACT

An invariant correlation digital system using a nonlinear filter is presented. The invariance to position, rotation and scale of the target is achieved via Fourier transform, mapping polar and Scale transform, respectively. We analyzed the performance of this filter with different nonlinearities $k$ values according to the peak-to-correlation energy (PCE) metric. We found experimentally the best $k$ value for rotation and scale and the confidence levels of the filters. The filter was applied to the complete alphabet letters where each letter is a problem image of 256x256 pixels in size. The results are presented and show a better performance when they are compared with linear filters.

Keywords: Invariant correlation, nonlinear filter, scale transform

1. INTRODUCTION

From the physiology standpoint of view, the three dimensional world perception through the sensorial visual system, in humans and some animals, is a basic action, but only unique in the human species based in the cognitive mental processes. The brain has the capability to recognize the form of the objects and the mechanism is governed by the same analytical functions used in optical and digital pattern recognition\footnote{1}, this means that between the retina and the visual cortex exists a complex logarithmic polar mapping. For the majority of the digital systems the process is very different, because of the great complexity of the retino-cortical and its natural composition in the sensorial visual system. The digital images are obtained in different environment conditions in an optical-digital sensor that causes several problems for the identification and characterization of the object using computerized vision systems and the study is limited to static objects, only bi-dimensional images are considered.

In this work, we obtain some results with new computational algorithms for the recognition of several objects, independently of its size, angular orientation and displacement. We choose the Scale transform\footnote{2}, instead of the well known Mellin-Fourier transform\footnote{3}, because is very suitable to scale change sensitivities. The nonlinear filter is introduced to realize the digital invariant correlation that give us information of the similarity between different objects. This kind of filter has advantages compared with the classical matched filter\footnote{4}, the only phase filter\footnote{5} and others linear filters, due its great capacity to discriminate objects, the maximum value of the correlation peak is well localized and the output plane is less noisy. We applied this method for the recognition of alphabetic letters in Arial style with one rotational degree and one percent increments for rotation and scaling, respectively, until a complete rotation of 360 degrees and 25% scale variation.

Further author information:

A.C.B.: acorbel@hotmail.com, phone: +52-(646)-1182095

J.A.B.: josue@cicese.mx, phone: +52-(646)-1750500 ext (25033)
2. METHODOLOGY
A nonlinear correlation filter denoted by NLF is expressed like

\[ NLF = |F(u,v)|^k e^{-j\phi(u,v)}, \quad 0 < k < 1 \]  

(1)

where \(|F(u,v)|\) is the absolute value of the Fourier transform of the object to be recognized, \(j = \sqrt{-1}\), \(k\) is the nonlinearity strength factor and \(\phi(u,v)\) is the phase of the Fourier transform.

The peak-to-correlation energy (PCE) performance metric is defined as

\[ PCE = \frac{|E[c(0,0)]|^2}{E[|c(x,y)|^2]}, \]  

(2)

where the numerator is the peak correlation intensity expected value and the denominator is the mean energy expected value in the correlation plane. We introduce the Scale transform due to its property of invariance to size and rotational changes.

The Scale transform and its inverse are given by

\[ D_f(c) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)e^{(-jc-\lambda)c} dt \]  

(3)

and

\[ f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} D_f(c)e^{jc+\lambda} dc. \]  

(4)

The Scale transform is a restriction of the Mellin transform on the vertical line \(p=-jc+1/2\), with \(c \in \mathbb{R}\). We used the 2-D Scale transform in polar coordinates \((r, \theta)\) with the log of the radial coordinate \(\lambda = \ln r\) that is expressed as

\[ D(c_\lambda,c_\theta) = \frac{1}{\sqrt{2\pi}} \int_{0}^{\pi} \int_{0}^{2\pi} \exp(-jc_\theta) f(\lambda, \theta) \exp[-j(\lambda c_\lambda + \theta c_\theta)] d\lambda d\theta. \]  

(5)

The steps to obtain the nonlinear filter is showed in Fig. 1a. First, we have the original image, \(f(x,y)\), consisting of the image of letter E (target). The next step is to calculate the Fast Fourier Transform (FFT), and according to the well known shift theorem\(^{10}\) which establish that a translation of the object in the spatial domain introduce a linear phase shift in the frequency domain, we obtain the modulus of the Fourier transform denoted as \(|F(w_x, w_y)|\), and of this way the displacement of the image is not affected in the Fourier plane. After this, we realized a high-pass filtering using a parabolic filter applied to the modulus Fourier transform. This kind of filter attenuates low frequencies while passing high frequencies that enhance sharp details, but cause a reduction in contrast in the image\(^{11}\). The next step is to introduce a scale factor given by \(\sqrt{r}\), where \(r\) is the radial spatial frequency, that is indispensable to differentiates the Scale transform from the Mellin transform. Next, Cartesian coordinates are mapped to polar coordinates to obtain the rotation invariance and a bilinear interpolation is done to the first data of the coordinate conversion to reduce the aliasing due to the log-polar sampling. A logarithmic scaling is made in the radial part in polar coordinates to realize the scale invariance. And taking the FFT in this step, we obtain the nonlinear filter.

Fig. 1b shows the steps for obtaining a composite nonlinear filter. In this case, there are a number of \(n\) images at the input of the system. The next step is to take the Fast Fourier Transform of each image and after doing this we have the total sum of all the FFT. The sum is then processed in the same way as was mentioned in Fig. 1a.
Fig. 1a. Flux diagram for obtaining the nonlinear filter

Fig. 1b. Flux diagram for obtaining the composite nonlinear filter
The steps to realize the invariant correlation with a nonlinear filter are shown in Fig. 2. In this figure, from the step (1) to the step (5), the procedure is the same as in Figs. 1a and 1b, except that now the steps (5) and (6) show the correlation procedure to obtain the invariant digital correlation, step (7), to position, rotation and scale using a nonlinear filter.

Fig. 2. Flux diagram representing the invariant correlation system with a nonlinear filter.

3. RESULTS AND DISCUSSION

The first task was to find by a numerical statistical experiment, the best $k$ nonlinearity strength factor value for the nonlinear filter. The results were box plotted, peak-to-correlation energy (PCE) vs $k$, one for the rotation and the other for the scale, by the mean value with two standards errors ($\pm 2*SE$) using the letter E in Arial style as the target. In both cases, the result for the maximum $k$ value was the same, $k=0.3$.

A box plot graph for the peak-to-correlation energy PCE vs the nonlinearities $k$ values is shown in Fig. 3 for the rotation case, from $k=0$ (phase only filter) to $k=1$ (matched filter) for the letter E Arial type used like the target and correlated with the 26 alphabet letters rotated, with one rotation degree increment, from 0 to 360 rotation degrees. This graph shows the mean value with one standard error ($\pm SE$) and two standard errors ($\pm 2*SE$) for the PCE. The number of images processed for each $k$ value was 9360. But there are 11 different $k$ values, that is, for $k=0, 0.1, 0.2, ..., 1.0$, so, we have a total of 102960 statistically processed images. This was done to determine the $k$ maximum value, that was $k=0.3$, by a numerical experiment.

From Fig. 3, we observe that for values $0.1 \leq k \leq 0.5$, the performance of the nonlinear filter is better than for the phase only filter. But for values $0.6 \leq k \leq 1$, the performance of the phase only filter was better than for the linear and nonlinear filter.
Fig. 3. Box plot graph for the peak-to-correlation energy PCE vs the nonlinearities k values, from k=0 to k=1 for the letter E type Arial correlated with the 26 alphabet letters rotated, with one rotation degree increment, from 0 to 360 rotation degrees.

Fig. 4. Box Plot graph for the peak-to-correlation energy PCE vs the nonlinearities k values, from k=0 to k=1 for the letter E type Arial correlated with the 26 alphabet letters scaled, with one percent scale increment, from -25% to +25%.
In a similar manner, a box plot graph for the peak-to-correlation energy PCE vs the nonlinearities k values is shown in Fig. 4 for the scale case, from \(k=0\) (phase only filter) to \(k=1\) (matched filter) for the letter E type Arial used like the target and correlated with the 26 alphabet letters scaled, with one percent increment, from -25% to +25%. This graph shows the mean value with one standard error (±SE) and two standard errors (±2*SE) for the PCE. The number of images processed for each k value was 1326. But there are 11 different k values, that is, for \(k=0, 0.1, 0.2, \ldots, 1.0\), so, we have a total of 14586 statistically processed images. This was done to determine the k maximum value, that was \(k=0.3\), by a numerical experiment.

From Fig. 4, we observe that for values \(4.0 \leq k \leq 1.0\), the performance of the nonlinear filter is better than for the phase only filter. But for values \(0.5 \leq k \leq 1.0\), the performance of the phase only filter was better than for the linear and nonlinear filter.

For both cases, the rotated and scaled, the result for the best k value was the same, \(k=0.3\). Both graphs look similar (Figs. 3 and 4).

The performance of the filter E in Arial style for rotation is shown in Fig. 5. The filter with a nonlinear factor \(k=0.3\) is correlated with each one of the 26 alphabet letter, with one rotation degree increment, from 0 to 360 degrees.

The performance of the filter E in Arial style for scale is shown in Fig. 6. The filter with a nonlinear factor \(k=0.3\) is correlated with each one 26 alphabet letters scaled from -25% to +25 with one percent increment.

From Fig. 6, we observe an overlap of the letter E with the F and L letters. This because the capacity of sensitivity to rotation and scale size has relation with the extension and structure of each character, for example, in this case E is more size-sensitive than an O but not sensitive for the F and L letters. Anyway to one SE the filter E recognize this character of the F letter. From rotation there is no overlap of the letter E as seen in Fig. 5. The filter E has showed a good discrimination between all the alphabet letters.

We have calculated the confidence level in percentage values for the nonlinear filters in the rotation and scaled cases. For the first case, Table 1 shows the results obtained for each one letter of the alphabet as a target, as is shown in the first row of the table, and correlated with the 26 alphabet letters, as is shown in the first column. These targets were rotated one by one degree increments, from zero to 360 rotation degrees and we used the nonlinear strength factor \(k=0.3\). The number of images processed for 26 filters correlated with the 26 alphabet letters, rotated one by one degree until complete 360 degrees, give us a total of 243360 statistically processed images.

This was done for determine the confidence level of each filter. We found only two overlapping between the filter and the character to be recognized and we denoted it as Not Considered (NC), due to its great similarity between them, the filter and the character, that was not capable to discriminate it. This occurs when there are high rotation-sensitive characters, as mentioned before due to its own characteristics of extension and structure.

For the scaled case, we calculated the confidence level in percentage values for the nonlinear filters applying the same method for the rotation case. Table 2 shows the results obtained for each one letter of the alphabet as a target, as is shown in the first row of the table, and correlated with the 26 alphabet letters, as is shown in the first column. These targets were scaled in increments of 1%, from -25% to +25% in scale, and we used the nonlinear strength factor \(k=0.3\). The number of images processed for 26 filters correlated with the 26 alphabet letters, scaled one by one percent until complete ±25%, give us a total of 34476 statistically processed images.

Of this way, we determined the confidence level of each filter. We found here 29 overlapping between the filter and the character to be recognized, and we denoted it as Not Considered (NC), due to its great similarity between them. But the results showed a better confidence levels for scale. This occurs when there are high scale-sensitive characters, as mentioned before due to its own characteristics of extension and structure.
Fig. 5. Performance of the filter $E$ in Arial style for rotation

Fig. 6. Performance of the filter $E$ in Arial style for scale
Table 1. Confidence levels in percentage for the filters, showed in the first row, correlated with the 26 alphabet letters, showed in the first column, rotated from 0 to 360 degrees with a nonlinear factor $\theta = 0.3$. The notation NC refers to Not Considered because there is an overlap between the filter and the letter.

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>L</th>
<th>M</th>
<th>N</th>
<th>O</th>
<th>P</th>
<th>Q</th>
<th>R</th>
<th>S</th>
<th>T</th>
<th>U</th>
<th>V</th>
<th>W</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>90</td>
<td>95</td>
<td>95</td>
<td>95</td>
<td>90</td>
<td>85</td>
<td>85</td>
<td>80</td>
<td>70</td>
<td>65</td>
<td>60</td>
<td>55</td>
<td>50</td>
<td>45</td>
<td>40</td>
<td>35</td>
<td>30</td>
<td>25</td>
<td>20</td>
<td>15</td>
<td>10</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2. Confidence levels in percentage for the filters, showed in the first row, correlated with the 26 alphabet letters, showed in the first column, scaled ± 25% with a nonlinear factor $\theta = 0.3$. The notation NC refers to Not Considered because there is an overlap between the filter and the letter.

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>L</th>
<th>M</th>
<th>N</th>
<th>O</th>
<th>P</th>
<th>Q</th>
<th>R</th>
<th>S</th>
<th>T</th>
<th>U</th>
<th>V</th>
<th>W</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>90</td>
<td>95</td>
<td>95</td>
<td>95</td>
<td>90</td>
<td>85</td>
<td>85</td>
<td>80</td>
<td>70</td>
<td>65</td>
<td>60</td>
<td>55</td>
<td>50</td>
<td>45</td>
<td>40</td>
<td>35</td>
<td>30</td>
<td>25</td>
<td>20</td>
<td>15</td>
<td>10</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Fig. 7. Scale invariant correlation using (a) a nonlinear filter with $k=0.3$ and PCE$=0.0759$, (b) a phase only filter with PCE$=0.0156$

Fig. 8. Rotation and scale invariant correlation using (a) nonlinear filter with $k=0.3$ and PCE$=0.0351$, (b) a phase only filter with PCE$=0.0141$
Fig. 9. Invariant correlation using (a) a nonlinear composite filter with $k=0.3$ and PCE=0.4704, (b) a composite phase only filter with PCE=0.1932

Fig. 10. Invariant correlation using (a) a nonlinear composite filter with $k=0.3$ and PCE=0.3879, (b) a composite phase only filter with PCE=0.1629
From the figures 7 to 10 we present some correlations examples using nonlinear filters, with $k=0.3$, that are compared with the phase only filters. The last four figures are for composite nonlinear and composite phase only filters. From these figures, we observe that for the nonlinear filter the correlation plane is less noisy and the correlation peaks are more defined and narrowed than for the phase only filter. In figures 7 and 8, the target and the problem images are the E letters in Times New Roman style, and from figures 9 and 10 the target and the image problems are letters in Arial style. In Fig. 7, the three peaks are along the scale axis that correspond to each E letter scaled in the problem image. Fig. 8 shows the scale and rotation in the problem image and the three peaks are over the scale axis and rotation axis that correspond to each of the letters of the problem image. In Fig. 9 we have a composite nonlinear filter and one peak centered on the correlation plane that indicates the recognition of $\text{EH}$ in the problem image. In Fig. 10, the letter $H$ is rotated 90 degrees in the problem image and there are two symmetrical peaks along the rotation axis on the correlation plane each corresponding to 90 and 270 degrees because the $H$ rotated 90 degrees has the same form that for 270 degrees. In all these cases, the PCE value for the nonlinear filter was better compared with the only phase filter.

4. CONCLUSIONS

The digital system for invariant correlation using nonlinear filter was tested with the maximum nonlinearity strength factor value $k=0.3$ that was determined experimentally for both scale and rotation. We used the whole alphabet letters in Arial style where each one of these letters was taken as a target and correlated with each one for the other letters to obtain the confidence level that we found by a rigorous statistical analysis for all the alphabet letters that proves our system work efficiently for the discrimination objects. This nonlinear invariance correlation method was applied using nonlinear, composite nonlinear, only phase and composite only phase filters for different objects scaled and rotated and we found a better PCE performance for the nonlinear filters.

REFERENCES