Detection of the Objects by the Cross-correlation

Moulay lhabib Elhadi
Département de mathématiques et informatique
Faculté des sciences Kenitra,
Université Ibn Tofail-Maroc

Mohammed Merzougui
Labo MATSI
EST-Oujda
Université Mohamed I

Moulay Youssef Hadi
Département de mathématiques et informatique
Faculté des sciences Kenitra,
Université Ibn Tofail-Maroc

Abdelaziz Mouloudi
Département de mathématiques et informatique
Faculté des sciences Kenitra,
Université Ibn Tofail-Maroc

ABSTRACT
The recognition of shape is a method aiming the identification of the motives from raw data in order to take a decision depending on the category assigned to this motive. The method of motive detection by

cross-correlation is a very efficient method in recognition of shape. Indeed, it permits to localize an object in an image. But, the value of the cross-correlation depends more of the level of gray of the image. In this article one proposes a solution of this inconvenience based on the Fourier Transform where by exploiting its phase for to localize an object while giving its position in the image.

Keywords
Recognition of shape, identification of motive, cross-correlation, Fourier Transform phase, localization.

1. INTRODUCTION
The recognition of shapes (or sometimes recognition of motives) means a set of techniques and methods aiming to identify some motives from raw data in order to take a decision depending on the category assigned to this motive. One considers that it's a branch of the artificial intelligence that calls extensively the techniques of automatic training and statistics [1].

The shapes or motives to recognize can be of very varied nature. It can be about visual content (code rod, face, fingerprint...) or resonant (recognition of speech), of medical images (X-ray, EEG, IRM...) or multi spectral (satellite images) and well of others.

However, to decipher a typed or handwritten text, to count some chromosomes, to recognize a tumor, a chariot or a plane of war, the understanding of the image, its classification always passes by the recognition of a shape. Several theoretical approaches have been developed [2].

Among these approaches, the one founded on thecross-correlation. In fact, this approach is used very often in recognition of shapes. But it maintains its dependence on the level of gray of the image.So, to eliminate this anomaly, one should use the phase of Fourier Transform.

In Section 2 , an introduction is given to the cross-correlation method. In section 3 presents a proposal for detecting objects in an image based on the Fourier transform while strengthening by the results of the application of the cross-correlation concept. At the end, a conclusion completes this article.

2. CROSS-CORRELATION
The concept of the cross-correlation aims to extract the rate of resemblance between two signals, the

one-dimensional signal (temporal signals for example) and the bi-dimensional (images). The mathematical definition of this concept, is that the cross-correlation between the two signals s(t) and s'(t), suppose one-dimensional in a first time, zero mean, is:

$$\text{xcorr}(s,s')(t)=\sum_{i=-N}^{N} s_i * s'_{i-t} \quad (2.1)$$

The philosophy of the cross-correlation is the search for the delay (or t translation) that allows to maximize the resemblance between s and s'. Indeed, it is only for a t delay that maximizes the resemblance that the sum takes place in a coherent with the product of s_i and s'_{i-t} all positive (in the order of s_i^2). In the contrary case , the product of the values of s_i makes itself with a signal s' of null middle value that has all odds to provide a result close to 0. So, t is a value sought to maximize the sum of the elements of the product s(i) with s'(i-t) when the t delay allows to s to look like best to s' shifted itself to t.

The principle of the cross-correlation consists of searching for, in segments of two images of a same object, the vectors of transfer that permit to recover the maximum of similarity between the two images.

This method remains inefficient because its value depends closely on the level of gray of the image [3][4].

3. PROPOSITION
The objective of this article is to define the frequency domain of an image, in other words, of its Fourier Transformation, of the describers permitting to recognize some objects in the images. The context is, therefore, the one of the treatment of images and the recognition of shapes.

In this article, one tries in particular, to take as a basis the Discrete Fourier Transform (DFT) . when it is desired to calculate the Fourier transform of a function x(t) using a computer, the computer having only a finite number of words of finite size, it is necessary to:

- discretize the time function,
- truncate the time function,
- discretize the frequency function.

$$X(f) = \int_{-\infty}^{+\infty} x(t) e^{-j2\pi ft} d(t) \quad (3.1)$$

Approaching the integral by a sum of duration of rectangles areas Te and limiting the integration time in the range [0, (N-1)Te] :

$$X(f) \approx T_e \sum_{n=0}^{(N-1)} X(nT_e)e^{-j2\pi fnTe} \quad (3.2)$$

Which gives for the frequency values fk = k fe/N

$$X(f) \approx T_e \sum_{n=0}^{(N-1)} X(nT_e)e^{-j2\pi\frac{nk}{N}feTe} \quad (3.4)$$

$$X(f) \approx T_e \sum_{n=0}^{(N-1)} X(nT_e)e^{-j2\pi\frac{nk}{N}} \quad (3.5)$$

It is not sophisticated approximation of X(f) , But it is widely used in practice under the name of DFT because it is an efficient algorithm called FFT (Fast Fourier Transform).

DFT is also used, when working with numerical sequences unrelated to a signal physical, to define a representation of the sequence on the basis of frequency functions.

Called discrete Fourier transform of a series of N terms x(0),x(1), . . . , x(N −1), Following N-term X(0),X(1), . . . ,X(N −1), defined by:

$$X(k) = \sum_{n=0}^{(N-1)} X(n)e^{-j2\pi\frac{nk}{N}} \quad (3.6)$$

In practice, N terms x(n) can be N samples of a sampled analog signal xn = x(nTe), and the N terms X(k) correspond to an approximation (within a multiplicative factor near T_e) of the transform this signal to Fourier N frequency points fk = fe / N, with k between 0 and N-1, ie f between 0 and fe

In the continuation, one is going to express the cross-correlaion (CC) by the formula of Fourrier Transform.

A classical template matching solution is the Sum of Squared Differences (SSD) - here with continuous functions :

$$SSDw = \int_w \left(f(t) - g(t)\right)^2 dt \quad (3.7)$$

SSD is related to CC :

$$\int_w \left(f(t) - g(t)\right)^2 dt$$

$$= \int_w (f(t)^2 + (g(t)^2 \; 2f(t)g(t)\; dt \quad (3.8)$$

$$= \int_w f(t)^2 dt + \int_w g(t)^2 dt - 2\int_w f(t)g(t)\; dt \quad (3.9)$$

If f is the pattern, the first term is constant.

Then:

$$\int_w \left(f(t) - g(t)\right)^2 dt$$

$$= Cte + \int_w g(t)^2 dt - 2\int_w f(t)g(t)dt \quad (3.10)$$

If the local energy of the image (g) is constant:

$$\int_w \left(f(t) - g(t)\right)^2 dt = Cte - 2\int_w f(t)g(t)dt \quad (3.11)$$

In this case, the SSD is the same as the CC

But the local energy of the image must be constant.

The Cross-Correlation (CC) is defined as (for real signals) :

$$(f * g)[n] = \sum_{k=-\infty}^{\infty} f[k]g[n+k] \quad (3.12)$$

$$(f * g)[n] = \sum_{k=-\infty}^{\infty} f[n-k]g[n] \quad (3.13)$$

CC is strongly related to convolution:

$$(f * g)[n] = f[-n] * g[n] \quad (3.14)$$

CC can also be easily expressed in Fourier Domain (fast computations) [5][6]:

$$FT[(f * g)][u] = \mathrm{F} * [\mathrm{u}].\mathrm{G}[\mathrm{u}] \quad (3.15)$$

So

$$(f * g) = \mathrm{FT}^{-1}[\mathrm{FT} * [\mathrm{f}]\mathrm{FT}[\mathrm{g}]] \quad (3.16)$$

4. APPLICATION

The idea is to make the image of pattern slip on the image of reference and for every position, to calculate the product of the pixels of the pattern with the pixels of the reference, then to make the sum of all these values. This operation is called the product of cross-correlation. The mathematical formulation is the following:

D(r,c)=−∞+∞I(k,l)P(r+k,c+l) (4.1)

Where I is the reference image, P is the pattern; and D the result. The local maximums of strong values of the cross-correlation correspond then to the positions where the pattern is present.

In practice, the cross-correlation is not nearly ever calculated under this shape. He/it is a lot more advantageous (in term of computer time) to do this operation in the domain of Fourier [6]:

D=FT−1(FT(I)·FT(P)∗) (4.2)

So the cross-correlation turns into a product pixel to pixel, which is fast. The implementation is direct enough, but it is necessary to pay attention to the following points,

- the Discrete Fourier Transformation is a function 2π periodic,
- the product pixel to pixel imposes that FT(I) and FT(P) are in the same size.

So the image of the pattern must be modified so that:

- its size is the same of the reference,
- the center of the image of the pattern corresponds to the origin of the axes.

5. RESULTS

5.1 Detection of the object (pattern)

Let the following image, considered as reference image (Figure1):

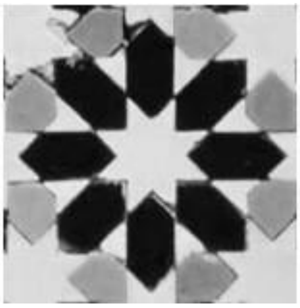

Fig 1: Image reference

One notices the pattern in the centre of the image. To find it, it is necessary to possess an image containing a copy of the object to search for first (the model), here it is called pattern (Figure2):

Fig 2: Image Pattern

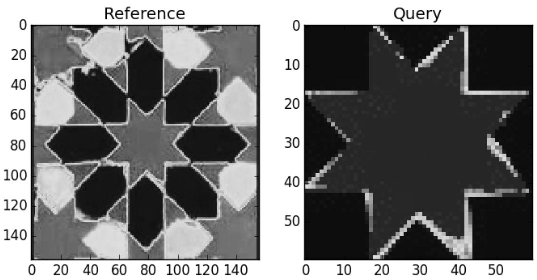

Fig 3: Images after change of the size

While applying the cross-correlation to the images of Figure3 one gets the following result:

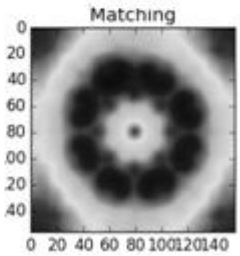

Fig 4: Result of motive detection

The motive being (more or less) in the center of the image, the shades of the center should be more important than the other to produce a local maximum of strong value.

One immediately notes that the local maxima don't correspond to the position of the object. In fact the only cross-correlation is not very efficient. Its value depends more on the gray level of its spatial structure.

Two solutions are proposed to palliate the problem of the sensitivity to the luminance:

- The normalized cross-correlation.
- Exploitation of the phase of the DFT.

The first solution to the problem of luminances sensitivity is to normalize the images before the comparison.

The used formula is the following [6]:

$$NCC_{I,P} = \frac{1}{N}\sum_{x,y}\frac{(I[x,y]-\bar{I})(P[x,y]-\bar{P})}{\sigma_I \sigma_P} \quad (5.1)$$

$\bar{I}$ is the average of I, σ_I is the standard deviation of I.

This solution is not landed for the following reasons:

- Very classic solution to template matching
- Same as for Pearson Correlation Coefficient (PCC)

In this work one is going to use the second solution that passes by detection of motives while exploiting the phase of the DFT.

Indeed, it's clear that (in theory) a cross-correlation could permit to localize an object in an image. The conclusion was that the cross-correlation is not very efficient because its value depends of the level gray of the images and rather little of their spatial information. Our suggestion consists of exploiting the properties of Fourier Transform, used to calculate the cross-correlation.

the formula used previously was:

$\widehat{D} = \hat{I} \cdot \widehat{P}^*$ (5.2)

$\hat{I}$ is the DFT of the I. The calculation consists therefore of taking the DFT of the image references I, the DFT of the image P pattern (in return for the described previously precautions) and to multiply pixel to pixel the two achieved images.

The Fourier Transform of a real image is an image whose values are the complex number. Let c a complex number, it can express itself in two ways [7]:

$$c = x + iy = A^{e\varphi} \quad (5.3)$$

where x , y , A and φ are real numbers.

x is the real part. y the imaginary part. A is the argument. φ is the phase

exploiting this last formula.

Since $\hat{I}$ is a complex image, it is also possible to write it under the shape of a module and an argument. The argument being called phase in general [8][9]:

$$\hat{I} = AIe^{\varphi I} \quad (5.4)$$

AI is then an image containing the module of every value of I, φI an image containing the phase.To understand what information is contained in AI and φI, it is necessary to isolate them from $\hat{I}$:

$$AI = |\hat{I}| \text{ et } \varphi I = \frac{\hat{I}}{|\hat{I}|} \quad (5.5)$$

Let's take an illustrative example. Here is the portrait of Joseph Fourier:

Fig 5: Portrait of Joseph Fourier

Here are the module and the phase of the DFT of this portrait:

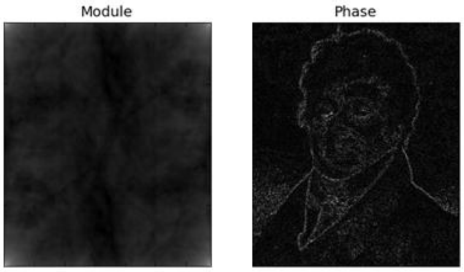

Fig 6: Module and phase of the DFT of portrait

It is then clear that the phase contains the contours of the image, therefore its structures.

The result is that :

To improve the detector of motive, it is therefore necessary to:

1. suppress the information of the module, because it contains some information on the level gray of the images.
2. keep the phase, because it contains the structures of the image.

The detector is modified to keep the phase of the images only:

$$D = FT^{-1}\left(\frac{\hat{I}}{|\hat{I}|} \cdot \frac{\widehat{P*}}{|\widehat{P*}|}\right) \qquad (5.6)$$

In the literature, this detection is known under the SPOMF denomination (Symmetric Phase Only Matched Filter) [10].

If one applies the new detector in search of motive to the images of the Figure 3, the result is much better.

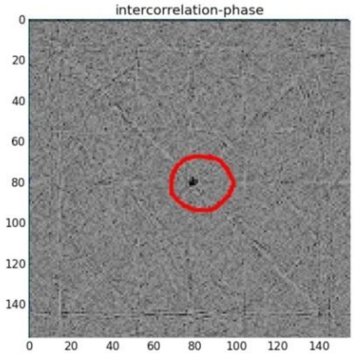

Fig 7: Result of motive detection while exploiting the phase of DFT

This time, the local maximum clearly distinguishes itself in the image and permits easy localization of motive. By against the first resulting output mounted in Figure 4 where the local maxima don't correspond to the position of the object which does not allow clearly locate the object. In this last result (fig 7) appeared the efficiency of new detector after removing the module information and the conservation phase.

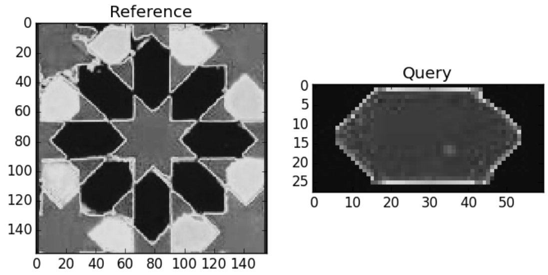

Fig 8: Images reference and pattern

A case where one has two processes of pattern:

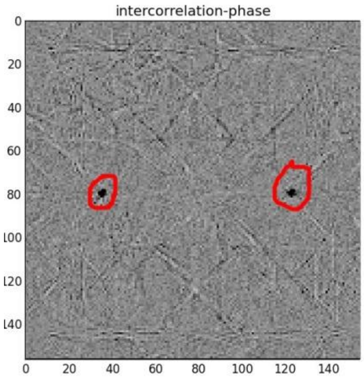

Fig 9: Result of motive detection with two processes of pattern

In this case one has two maximal premises that distinguish themselves clearly in the image. But the other processes, whose orientation is different to the one of the image pattern, are not detected.

So it is necessary to improve the detector again to make it robust to the orientation problem.

5.2 Location of the pattern in the image

In the example where one has only one instance of pattern, there is only one position to find. One looks for the maximum position in the image.

The found position is (80, 79), which corresponds well to the position of pattern in the image.

To improve the precision of localization it is necessary to use the methods of linear interpolation such as spline, quadratic ...

6. CONCLUSION

This work shows the ineffectiveness of the cross-correlation only to detect objects in an image. In fact, its value depends more on the gray level of its spatial structure. The new method based on the use of the phase of the Fourier transform of the image provides a good pattern of localization. However, in case the orientation and sizes of instances are different from the image pattern, or if the variations of orientation and scale are weak, the method cannot function anymore.

In future work, it is necessary to improve this method to finally develop a detector that more resistant to these variations.

7. REFERENCES

[1] Richard O. Duda, Peter E. Hart, David G. Stork, Pattern classification, Wiley-interscience, 2001.

[2] Dietrich Paulus and Joachim Hornegger (1998) Applied Pattern Recognition (2nd edition), Vieweg.

[3] J.-M Friedt Auto et intercorr_elation, recherche de ressemblance dans les signaux : application a l'identifcation d'images floutées, 29 avril 2011

[4] F. Roddier, Distributions and transformation de Fourier, Edisciences International, Paris (1978) [chapitre 11].

[5] D.W. Kammler, A first course in Fourier analysis, Cambridge University Press, 2008.

[6] Jean Dhombres et Jean-Bernard Robert, Fourier, créateur de la physique mathématique, collection "Un savant, une époque", Belin, 1998

[7] Dominique Flament, Histoire des nombres complexes : Entre algèbre et géométrie, Paris, CNRS Éditions.

[8] Van des Schaaf A., Van Hateren J., "Modelling the Power Spectra of Natural Images : Statistics and Information",Vision Research, vol. 36, n°17, p. 2759-2770, 1996.

[9] Q. Chen, M. Defrise and F. Deconinck, "Symmetric phase-only matched filtering of Fourier-Mellin transforms forimage registration and recognition", IEEE pattern analysis and machine intelligence, vol. 16, 1994, p. 1156-1168.

[10] Qin-sheng Chen, Michel Defrise, and F. Deconinck Member, IEEE. Symmetric Phase-Only Matched Filtering of Fourier-Mellin Transfoms For Image Registration and Recongition. Vol 16,NO.12 DECEBMER 1994.