Vectorial Signatures for Pattern Recognition

Jesús Ramón Lerma-Aragón\textsuperscript{1} and Josué Álvarez-Borrego\textsuperscript{2}

\textsuperscript{1}Facultad de Ciencias, Universidad Autónoma de Baja California, México
\textsuperscript{2}CICESE, División de Física Aplicada, Departamento de Óptica, México

1. Introduction

Due to the variety of shapes and sizes that present both living organisms and static objects, the necessity to look for automated systems of identification, both in industry and scientific research has arisen. During the 1960s, the scientific community in the field of optics has used the Fourier transform and other types of mathematical transformations for pattern recognition, taking advantage of their different properties; for example, invariance to position, rotation and scale.

Many invariant descriptors use the Fourier transform to extract invariant features. The Fourier transform has been a powerful tool for pattern recognition. One important property of the Fourier transform is that a shift in the time domain causes no change in the magnitude spectrum. This can be used to extract invariant features in pattern recognition.

In the last few years, they have been used as a tool in the digital processing of pattern recognition. Recently several books have been published (Pratt, 2007, Gonzalez et al., 2008 and 2009, Cheriet et al., 2007, and Obinata, 2007), that show us a general view, the progress and development of pattern recognition, as well as different tools for image pre-processing, extraction, selection and creation of features, classification methods, using different types of transforms, etc. However, some studies have focused on solving the problem of invariance to rotation and scale (Cohen, 1993, Casasent, 1976a, Pech et al., 2001 and Pech et al, 2003, Solorza and Álvarez-Borrego, 2010, Solorza and Álvarez-Borrego, 2011).

For practical implementation of the theory it should be considered, according to Casasent and Psaltis (1976a, 1976b, 1976c), a system that is invariant to scale through the manipulation of the Fourier transform directly changing the input function. We must also consider the practical realization of the Mellin transformation given by a logarithmic mapping of the input stage followed by a Fourier transform.

Schwartz (1994) found that there is strong evidence in many physiological and psychophysical visual systems (including human) use of such logarithmic mapping between the retina and visual cortex. The motion of changing a scaling to a shift by a logarithmic transformation of the coordinates occurs in many areas. In fact, the log-polar coordinate system described above seems to have a biological analogue. Since biological systems can shift objects to the centre of the field of view by movement of the eyes, it seems logical that a mapping that facilitates scale and rotation invariant recognition would be most useful. This is not to say that biological systems also compute Fourier transforms of these representations.
Other methods can take advantage of representation in which transformations are reduced to shifts.

The different techniques used today have difficulties, the main one being that the calculations involve a high computational complexity. Moreover, in this case, the complexity of the operation performed is directly proportional to the resolution of the images used. The computational cost associated inevitably increase the time required for comparison. Emerging applications of computer vision and pattern recognition require the development of new algorithms. The vectorial signatures techniques have an important role to play in this context, given their simplicity and low computational requirements.

In this chapter we will focus on examining the capabilities of vectorial signatures, as a method of recognizing objects, and we will consider inside this algorithm invariance to position, rotation and scale. It will be applied here to the recognition of alphabetical letters in Arial typeface, copepod species and some butterflies species (color images). However, this algorithm can be applied to any kind of object. The system was developed by taking advantage of the properties of the Fourier, Mellin and Scale transforms.

2. Mathematical description

In this section we present some basic concepts, such as the definition of Fourier, Mellin and Scale transforms, tools that are of great usefulness in audio and image processing, noise reduction signals (such as white noise) frequency analysis of any discrete signal, materials analysis and statistical synthesis by inverse Fourier transform, and so on.

2.1 Fourier transform

The Fourier transform $F(u)$ of a continuous function $f(x)$ of a single variable is defined

$$F(u) = \int_{-\infty}^{\infty} f(x)e^{-j2\pi ux} \, dx$$

Where $j = \sqrt{-1}$. Consequently, the inverse Fourier transform (IFT) of $F(u)$ results in the original function, described by

$$f(x) = \int_{-\infty}^{\infty} F(u)e^{j2\pi ux} \, du$$

2.1.1 The Fourier-Mellin transform

Mellin transform is especially useful for the scale invariant (Bracewell, 1978). This has been applied to the spatially varying image restoration and in the analysis of networks that vary with time, among others.

The two dimensional Mellin transform $M(ju, jv)$ of a function $f(x, y)$ along the imaginary axis is defined by Casasent and Psaltis (1976a)

$$M(ju, jv) = \iint_{0}^{\infty} f(x, y)x^{-ju-1}y^{-jv-1} \, dx \, dy$$
2.1.2 The scale transform

Cohen (1993) introduced the “scale transform.” This transform is said to be scale invariant, thus meaning that the signals differing just by a scale transformation (compression or expansion with energy preservation) have the same transform magnitude distribution. Cohen showed that the scale transform is a restriction of the Mellin transform on the vertical line $p = -jc + 1/2$, with $c \in \mathbb{R}$.

We choose the scale transform, which is a restriction of the Fourier-Mellin transform (Casasent, 1976c). The main property of the scale transform is the scale invariant property.

If we call $c$ the scale variable, then in two dimensions the scale transform and its inverse is given by Cristobal & Cohen (1996).

$$D(c_x, c_y) = \frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} \int_{0}^{\infty} f(x, y) e^{-jc_x \ln x - jc_y \ln y} \frac{dx dy}{\sqrt{xy}}, \quad (4)$$

$$f(x, y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} D(c_x, c_y) e^{jc_x \ln x + jc_y \ln y} \frac{dc_x dc_y}{\sqrt{xy}}, \quad (5)$$

3. Implementation

In this section we will use two-dimensional scale transformations now written in polar coordinates $D(c_r, c_\theta)$, because this transform is invariant to changes in size and rotation

$$D(c_r, c_\theta) = \frac{1}{\sqrt{2\pi}} \int_{0}^{2\pi} \int_{0}^{2\pi} f(r, \theta) r^{-jc_r} \frac{dr d\theta}{\sqrt{r}}, \quad (6)$$

and taking the logarithm of the radial coordinate, that is $\lambda = \ln r$, we get

$$D(c_\lambda, c_\theta) = \frac{1}{\sqrt{2\pi}} \int_{0}^{2\pi} \int_{0}^{2\pi} f(\lambda, \theta) r^{-jc_\lambda} \frac{d\lambda d\theta}{\sqrt{\lambda}}, \quad (7)$$

Applying the transformation described above, we built a new algorithm. Fig. 1 shows the block diagram of the methodology used. The target ($I_i$) to be recognized is denoted by the function $f(x, y)$ (step 1) and the modulus of its Fourier transform ($F$) is obtained (step 2). A parabolic filter to the modulus of $F$ is applied (step 3). In this way, low frequencies are attenuated and high frequencies are enhanced in proportion to $(w_x)^2$, $(w_y)^2$ (Pech-Pacheco et al., 2003), where $w_x$ and $w_y$ are the angular frequencies. The scale factor $\sqrt{r}$, (step 4) is applied to the results of step 3 ($r$ is the radial spatial frequency). This process is what differentiates the scale transform from the Mellin transform. After these steps, we mapped the Cartesian coordinates to polar-logarithmic coordinates (step 5), to obtain invariance to rotation. In this step we introduced a bilinear interpolation of the first data of coordinate conversion. This is done to avoid the aliasing due to the log-polar sampling. A subimage denoted by $M(\lambda, \theta)$ (Lerma-Aragón & Álvarez-Borrego 2008) is chosen from step 5, which contains more information of the target and a bilinear interpolation is done again in order to resize the subimage data to its original size (step 6). From the subimage, we obtain two 1-D
vectors by projecting $M(\lambda, \theta)$ onto the $x$ and $y$ axes (steps 7 and 8). In other words, we compute the marginals, according to the equations

$$M(\lambda_m) = \sum_{k=1}^{n} M(\lambda_m, \theta_k),$$  \hspace{1cm} (8)$$

$$M(\theta_n) = \sum_{k=1}^{n} M(\lambda_k, \theta_n).$$  \hspace{1cm} (9)$$

Fig. 1. Block diagram of used procedure

A modulus of the Fourier transform is calculated in order to obtain the rotation and scale signatures (equations 10 and 11), which will be the two unidimensional vectors for the target (steps 9 and 10).

$$V_1(w_\lambda) = |F[M(\lambda_m)]|,$$  \hspace{1cm} (10)$$

$$V_2(w_\theta) = |F[M(\theta_n)]|.$$  \hspace{1cm} (11)$$

To evaluate this algorithm, the target image ($I_i$) and an image problem ($I_j$) ($I_j$ is denoted by the function $g(x,y)$ either rotated or scaled) are selected, to which we apply the procedure described above. To determine the similarity between the images $I_i$ and $I_j$ in the database, using the Euclidean distance ($E_d$) is calculated among their average firms using the following equation.
Thus, the signatures of all the $j$ images of the database can be compared with any image $i$ to be recognized.

### 3.1 Application of the algorithm

To apply the algorithm based on the theory described above, consider primarily a target image, in this particular case, we consider a simple image (rectangle) of size 256 X 256 pixels and black background, as shown in Fig. 2(a), which is calculated the magnitude of the Fourier transform, which can be seen in Fig. 2(b). We used Matlab colormap command to change to a colour figure and of this way clearly visualize each step of the procedures. When these values are scaled linearly for display, the brightest pixels will dominate the display, at the expense of lower (and just as important) values of the spectrum. If, instead of displaying the values in this manner, we apply logarithmic transformations (Gonzalez, 2009) to the spectrum values, then the range of values of the result is a more manageable number. Fig. 3 shows the result of scaling this new range linearly and displaying the spectrum. The wealth of detail visible in this image compared to a straight display of the spectrum is evident from these pictures.

$$E_d = \sqrt{\sum \left[ V_{1f}(w_\lambda) - V_{1g}(w_\lambda) \right]^2 + \left[ V_{2f}(w_\theta) - V_{2g}(w_\theta) \right]^2}.$$  \hspace{1cm} (12)
A parabolic filter (Fig. 4) to the modulus of $F$ is applied, in this way, low frequencies are attenuated and high frequencies are enhanced as shown in Fig. 5(a). The scale factor $\sqrt{r}$ is applied to the previous results (Fig. 5b). A subimage is chosen from Fig. 6(a) corresponding to the zone that contains the major quantity of information. In the Fig. 6(b) results are observed.

From the subimage $M(\lambda, \theta)$ seen in Fig. 6(b), we compute the marginal functions, using the following equations

$$g(x) = \sum_{\theta} M(\lambda, \theta),$$  \hspace{1cm} (13)

$$h(x) = \sum_{\lambda} M(\lambda, \theta).$$  \hspace{1cm} (14)

In the Fig. 7(a) the subimage (Fig. 6b) appears, in three dimensions, and in 7(b) and 7(c) the marginal functions can be observed, with which scale and rotation vectors are obtained.
Finally, Fig. 8 shows the modulus of the Fourier transform of the vectors showed in the Fig. 7(b) and 7(c) respectively, which defines the vectorial signatures of the image, and they are used for comparison with respect to another image.

Fig. 6. (a) Image 5(b) in polar-logarithmic coordinates. (b) Selected subimage

Fig. 7. (a) Tridimensional image of Fig. 6(b). (b) Vector of the summation on the rotation axis. (c) Vector of the summation on the scale axis

4. Computer simulations in grayscale images

To evaluate the performance of the vectorial signatures, images of different types were used as input images.
4.1 Simulation using letters

The first case analyzed corresponds to images of the different letters of the alphabet. Each letter is an image of 256 X 256 pixels of black background with a centered white Arial letter size 72. The methodology described above was applied to these letters. The behavior of the Euclidean distance ($E_d$) for the target E versus itself with respect to rotation from 0 to 359º with variations of one degree was studied. Because of the letter distortions, when is rotated, due to the square pixels, the curve representing the $E_d$ is cyclic and with symmetry of 180º (Fig. 9).
Fig. 10. Euclidean distance of the letter E compared with other letters

Fig. 11. Statistical distance behavior
To evaluate the distinguish ability between the 26 letters of the alphabet; each one was rotated 360 degrees, with a variation of one degree. Simulation was performed to determine the difference between them, using the 9360 images (26x360). In Figure 10 shows the behavior of the letters that have greater similarity with the reference (E); the result shows a very clear separation between the values of the Euclidean distance of the letter E compared with the rest of the letters, which allows being able to identify it. When an image is similar to another one, the $E_d$ has a minimum value. Statistic was performed and the mean value ± 2SE (two standard error) was calculated. This algorithm has at least a 95.4% level of confidence for this case (Fig. 11 shows the comparison with all the letters).

To analyze the behavior to changes in scale, changes were tested by 90% to 110% with respect to the target image, with increments of 0.5%. Fig. 12. shows the results obtained by reference to the letter E and compared again with all the letters of the alphabet.

4.2 Simulation using copepod species

Copepods are the dominant group in marine zooplankton; they constitute at least 70% of the planktonic fauna. This group of organisms has a great diversity. There are about 11,500 species of copepods described by Humes (1994) and the number is steadily increasing with the description of new species, such as those from anchialine caves (Fosshagen, 1991), hydrothermal vents (Humes, 1991), and those previously reported under the name of another species (Bradford, 1976; Soh & Suh, 2000).

Copepods are of prime importance in marine ecosystems. The majority of copepods feed on phytoplankton, forming a direct link between primary production and commercially...
important fish, such as sardine, herring, and pilchard. Copepods are also the main food source for a great variety of invertebrates. Studies on copepod abundance and species composition are particularly relevant, because most larvae of commercial fish feed on copepods. Hence, changes in the abundance of these plankters from year to year, may determine interannual population fluctuations of the commercially exploited fish stocks in a particular region.

For this study, adult stages of different copepod species were separated from several plankton samples. The specimens were observed with an optical microscope and their images were digitally captured using a charge coupled device camera (CCD). We used 14 images from seven different species of copepods each with male and female samples (Fig. 13). Since the background noise in all the images is mostly repetitive, the images where cleaned using some methods described in Guerrero & Álvarez-Borrego (2009). Fig. 14 shows the original and its pre-processed image for a female specimen *Rhincalanus nasutus*.

![Fig. 13. Species of copepod used in the study: (a) Calanus pacificus female; (b) C. pacificus male; (c) Rhincalanus nasutus female; (d) R. nasutus male; (e) Centropages furcatus female; (f) C. furcatus male; (g) Pleuromamma gracilis female; (h) P. gracilis male; (i) Temora discaudata female; (j) T. discaudata male; (k) Acartia tonsa female; (l) A. tonsa male; (m) Centropages hamatus female; (n) C. hamatus male](image-url)

The image of one of the calanoides *Calanus pacificus* female Fig. 13(a) was used as a target to discriminate between copepod species and sex from the rest of the organisms, the image size
Fourier Transform – Signal Processing

is 256 X 256 pixels of black background with a centered copepod. The image was rotated 360° in increments of 1° to recognize a target in an input scene; signatures were compared using the Euclidean distance ($E_d$) between the target and the input image and, because there is no variation in the size of copepod adults, we did not analyze the changes of scale.

In order to see if this methodology has a good performance for recognizing the target image when compared with other images that don’t correspond to the copepod chosen as reference, numerical calculations were performed for the entire set of 14 images showed in Fig. 13.

![Image](image1.png) ![Image](image2.png)

Fig. 14. (a) Original image. (b) Cleaned images

Fig. 15 shows the performance of $E_d$ for the fourteen pictures rotated 360 degrees. The Fig. 13(a) was used as a target, corresponding to a female *Calanus pacificus*. We can see a very good separation of the $E_d$ for the target with respect to the calculation of the $E_d$ for the non-target images. As can be seen in Fig. 15, the curve of the copepod (Fig. 13a) represents an $E_d$ with minimum values, which distinguish it from the rest of the copepods. Statistic was realized and the mean value ±2 SE was calculated. We can see this algorithm has at least a 95.4% level of confidence for this case (Fig. 16).

![Graph](graph.png)

Fig. 15. Change the angle of rotation against $E_d$
4.3 Average signature

Like a second set of test images 30 females and 30 males from each of the 7 different species of copepods were used, thus forming a database of 420 images. To have an identification system that contains information from several copepods, vector composed signatures were designed, which is derived from adding the signatures of 10 different copepods and to get its average, which is now the reference vectorial signature (Fig. 17).

In Fig. 18 is shown the application of this method, we can see an excellent separation of the Euclidean distance mean, which corresponds to Centropages furcatus female (Group e), used like target, with respect to the calculation of the $E_d$ from the rest of the groups. Statistic was realized and the mean value ± 2SE was calculated. We can identify the species and sex of the target with the confidence level was about 95.4%.

4.4 Comparison with other algorithms

To evaluate the performance of the algorithm, we compare it with respect to that published by Álvarez-Borrego and Castro-Longoria, considering that both are similar in application to pattern recognition and methodology based on the properties of the Fourier transform, in the second method is pattern matching using the correlation.

Comparing the results of the two algorithms applied to the same sets of images, we observed that in the second is not possible to distinguish the selected image as reference, with respect to the others. Regarding the computational cost, the processing time of our procedure is approximately 20% lower.
Fig. 17. Average vectorial signature

Fig. 18. Statistical behavior of the Euclidean distance, in the case of vectorial signatures composite of 10 images of Centropages furcatus female (Group e)
5. Computer simulations in color images

One of the most important characteristics, which account for visual pattern recognition is the color, and has been used in a wide range of applications in different areas of knowledge, such as evaluation and identification of textiles, flowers, microscopic images, surface corrosion, species recognition, and so on. In our daily life, our vision and actions are influenced by a variety of shapes and colors. The introduction of color increases the amount of information in pattern recognition, so that discrimination can be improved considerably.

Color is a perceived phenomenon and not a physical dimension like length or temperature. A suitable form of representation must be found for storing, displaying, and processing color images. This representation must be well suited to the mathematical demands of a color image processing algorithm, to the technical conditions of a camera, printer, or monitor, and to human color perception as well. These various demands cannot be met equally well simultaneously. For this reason, different representations are used in color image processing according to the goal. For these use different models that are known as color space (Gonzalez & Woods, 2008, Kosch & Abidi, 2008, Westland & Ripamonti, 2004) were used.

One of the most used color space is RGB, this model is based on the additive mixture of the three primary colors: red (R), green (G) and blue (B). When dealing with color images, there are several concepts that are inherently quite different. For example, if we treat the RGB signals at a pixel as a three-dimensional vector, a color image becomes a vector field, while a monochromatic image is a scalar field.

5.1 Methodology

Fig. 19. shows the block diagram of the methodology applied to RGB color spaces: in the first step, is select the color image to be used like a target ($I_i$) which is denoted by $f_i(x, y)$, where $i = R, G, B$, then each image is decomposed into its respective RGB channels [$f_R(x, y), f_G(x, y), f_B(x, y)$]. To have an identification system that contains information from the components $f_i(x, y)$, a vector-composed signatures were obtained, which is now the average vectorial signature $V_1(w_\lambda)$ and $V_2(w_\theta)$, for the test image ($I_j$) represented by $g_j(x, y)$ is performed the same process where were obtained $V_1g(w_\lambda)$ and $V_2g(w_\theta)$. To determine the similarity between the two images $I_i$ and $I_j$, the value of the $E_d$ is calculated.

5.2 Simulation for butterflies

To test this algorithm a set of images shown in Fig. 20 was used. With different types of butterflies, each image has a size of 256 x 256 pixels. In addition it is estimated that the number of species of butterflies in the world varies between 15,000 and 20,000, which according to the North American Butterfly Association (NABA), about 2000 are located in México.

In Fig. 21 we can see the statistical behavior of the Euclidean distance for the case of rotation of the image 360 degrees in one degree increments, using as reference the butterfly (10), which corresponds to a specimen Archceoprepona amphimachus amphiktion, which at first glance is very similar to others butterflies, for example with butterfly (11) belongs to same genus but different species. As expected, the minimum value of the distance corresponds to the target. We can see that this algorithm has a very good performance in to recognize the target when is compared with the others butterflies. Numerical simulations were performed...
considering as a reference each of the 18 butterflies and in all cases could be distinguished from the rest.

Fig. 19. Block diagram of used procedure

To analyze the invariance to scale changes were made from 70% to 130%, with variations of one percent each of the test images. The analysis of the statistical behavior of the distance is shown in Fig. 22, for the case where the butterfly (3) is used like target. The results indicate that there is separation of the reference image with the rest of the butterflies, which can be identified without difficulty, regardless of changes of scale.

Fig. 21. Statistical behavior of the $E_d$, for an Archaeoprepona amphimachus amphiktion specimen

Fig. 22. Statistical behavior of the $E_d$, for a Agraulis vanillae incarnata specimen
6. Conclusion

In this chapter it was developed a system for pattern recognition with the utilization of vectorial signatures, based on the well-known relation between scale and Fourier transform, and has been performed to be practical and accurate. To test the built model, simulations were made and the results were analyzed. Conducting various numerical tests with letters of the alphabet, with variations in the scale and rotation, the results obtained from a comparison of the Euclidean distances show that it is possible for identification and discrimination with a high confidence level (at least 95.4%). The system was applied to the recognition of copepod species and sex. With vectorial signatures it is possible to identify, with a confidence level of 95.4%, the species and sex of various species of copepods. A modification of the algorithm in color image recognition, using average vector signature from the information contained in the RGB channels was done. For this case images of real butterflies were used, finding that the system can identify the species you choose as a reference, when compared with images of different butterflies, for any angle of rotation that contains, as well as changes in the scale of 70% to 130%.

This work contributes to increase the potential recognition systems for use in applications requiring assessment and interpretation of data, applications that were traditionally performed by human vision or specialists trained technicians and are still far from automation.

7. Acknowledgment

This document is based on work partially supported by UABC and CONACYT under Project grants 102007 and 169174.

8. References


