

## Traitement, Synthèse, Technologie et Applications

BIARRITZ - Mai 1984 -

ANALYSE DE SCENES PAR DES OPERATEURS FORME-INVARIANTS

SCENE ANALYSIS BY MEANS OF FORM-INVARIANT OPERATORS

C.Braccini, G.Gambardella, A.Grattarola

DIST - Dipartimento di Informatica, Sistemistica e Telematica  
Università di Genova, Viale F.Causa 13, 16145 Genova, ITALY**RESUME**

Le filtrage forme-invariant est introduit et les correspondents opérateurs sont dérivés. On examine aussi les réalisations et les applications à la reconnaissance des formes indépendante des dimensions. Enfin, on présente des exemples de ces filtrages pour l'analyse des scènes.

## 1. INTRODUCTION

Form-invariant (or scale-independent) operators are needed in several image processing applications, such as computer vision and robotics. Here, in fact, a 3-D scene must be analyzed independently from the distance of the acquisition system, i.e. from the scale of the resulting 2-D image (or images) /1,2/.

In order to achieve some higher-level description of the scene, two preliminary stages of analysis are usually performed:

- i) some preprocessing, aimed, e.g., to reduce the noise or enhancing the contours;
- ii) the recognition of meaningful features in the image, such as edges or shapes or even objects.

This latter stage of analysis can be implemented by means of various scale-independent recognition methods that are based either on linguistic techniques or on some suitably developed iconic matching algorithms. In any case the global processing represented by the two above stages turns out to be scale-invariant if and only if the preprocessing operators are also scale-invariant.

This paper is focused on scale-invariant signal processing techniques that can be applied to both preprocessing and recognition problems.

The general problem of linear scale-invariant filtering is firstly investigated and its general solution is presented. It is shown that the linear scale-invariant filters are a particular class of the linear shift-variant filters. Several particular cases

**SUMMARY**

Scale-invariant signal processing is first defined and the corresponding operators are derived. Their implementation schemes are discussed as well as their use in size-independent pattern recognition. Some experimental results of such a filtering are presented in the framework of a particular scene analysis system.

of this class are then discussed in relationship to their processing properties (such as point or line symmetries) as well as to their implementation. This latter can be performed either via the compatible transform, i.e. the Mellin transform, or through a classical shift-invariant processing in a suitable computational space.

The application of scale-invariant operators to scene analysis is then discussed from the viewpoint of its effects on the iconic matching of scaled (and rotated) patterns: some experimental results of this kind of filtering are also shown, in the context of a particular scale-invariant recognition algorithm.

## 2. FORM-INVARIANT OPERATORS AND THEIR IMPLEMENTATIONS

The particular kind of processing that we call form-invariant or scale-invariant is characterized by the property that it translates scale changes of the input signal into scale changes of the output, so that no distortion effects take place. Mathematically, if  $f(x,y)$  and  $g(x,y)$  are an input-output pair for such processing, the output corresponding to the scaled input  $f(a_1x, a_2y)$  must be  $g'(x, y) = \alpha g(b_1x, b_2y)$  with  $\alpha$ ,  $b_1$  and  $b_2$  functions of  $a_1$  and  $a_2$ . If we consider only the linear systems, it has been shown /3/ that the most general class of operators satisfying the above property is defined by the set of weighting functions or kernels of the kind



$$w(\xi, \eta; x, y) = y^{-\alpha} w_0(\xi/x^{\sigma_1}, \eta/y^{\sigma_2}; x/y^{\nu}) \quad (1)$$

where  $w_0$  is any regular function of its variables,  $(\xi, \eta)$  and  $(x, y)$  are the input and output domains, respectively, and the other parameters are real constants, affecting the output scaling factors  $\alpha$ ,  $b_1$  and  $b_2$ . From eq. (1) it is apparent that linear form-invariant operators can be found only within shift-variant systems.

Several subclasses of form-invariant operators defined by eq. (1) are of particular interest from the viewpoints of applications or implementation. For example, the shift-variant input-output relationship based on the kernels of eq. (1) becomes a simple algebraic relationship (similar to the one obtained for the shift-invariant systems in the Fourier domain) in a suitably transformed domain when the weighting functions belong to the subclass

$$w(\xi, \eta; x, y) = y^{-\alpha} w_0(x/\xi^{\tau_1}, \eta/\eta^{\tau_2}) \quad (2)$$

In fact, for such a subclass the compatible transform exists and turns out to be the Mellin transform, defined as

$$\tilde{F}(s_1, s_2) = \iint_{\infty} f(x, y) x^{s_1-1} y^{s_2-1} dx dy \quad (3)$$

The implementation of the form-invariant operators of eq. (2) can therefore be performed by multiplying the Mellin transforms of the input image and of the kernel. Such transforms can be computed, in turn, as Fourier transforms of the original functions evaluated along logarithmically warped space axes.

Another subclass of eq. (1), exhibiting particular interest in several applications, is the one characterized by the circular symmetry of the weighting functions and the linear expansion (or contraction) of their size with the distance from a reference point, or origin, in the  $(x, y)$  plane (e.g. the geometric center of the image). Such operators are of the form

$$w(\xi, \eta; x, y) = w_c(\rho/r, \theta - \varphi) \quad (4)$$

where  $(r, \theta)$  and  $(\rho, \varphi)$  are the polar coordinates of the  $(x, y)$  and  $(\xi, \eta)$  planes, respectively. They are encountered in applications like image reconstruction from projections /4/, space-variant optical distortions /5/, peripheral visual system models /6/, and are the ones used in the examples shown in this paper. The above mentioned characteristics of these operators can be exploited to obtain a simpler implementation scheme with respect to the class of eq. (2). In fact, if the space domain input-output relationship is written using polar coordinates, a convolution comes out which is space-variant along the radial axis and space-invariant along the angular coordinate. It turns out, therefore,

that the compatible transform is now the mixed Mellin-Fourier transform with respect to the  $r$  and  $\theta$  variables, respectively. The corresponding implementation complexity is now reduced since only 1-D Mellin transforms are to be computed.

The problems raised by the above mapping from cartesian to polar coordinates are now examined, while discussing a different approach to the implementation of the subclasses defined in eq.s (2) and (4). They can in fact be reduced to the cascade of suitably chosen shift-invariant operators and coordinate mappings. These latter are easily shown to be the logarithmic ones: along both the cartesian axes in the case of eq. (2) and along the radial axis only in the case of eq. (4). It is worth noticing that, if the shift-invariant systems are implemented through the Fourier transform, the overall processing resulting from the coordinate mapping and the Fourier transform is equivalent to computing the Mellin transform.

Two comments are in order on the above introduced coordinate mappings. The first one refers to the properties of the mapped plane for pattern recognition and scene analysis applications. In particular, the so called logarithmic conformal mapping defines a computational space of coordinates axes  $(\ln r, \theta)$  where image description, processing and recognition schemes can be designed that are invariant under scene viewing changes /2,7,8/. This property is typical of the human visual system, where such a mapping seems to take place. The second comment concerns efficient ways of implementing the mapping, that requires a number of operations of the order of  $N^2$  (for a  $N \times N$  image) and a non-structured access to the data, and therefore represents a severe limitation, e.g. for real time (TV rate) applications on the usual raster scan image processing architectures. One possibility that we are exploring is based on a multi-processor system, where each processor is devoted to processing a sector (or a portion of a circular strip) of the image, built by storing a part of each scan line, as schematically indicated in Fig. 1. In this way both the input and output data are accessed on a line-by-line basis and the overall achievable speed is a function of the number of processors.

### 3. SCALE-INVARIANT PROCESSING IN SCENE ANALYSIS

In this section we discuss the use of form-invariant operators in the first stages of a computer vision system and compare the results with those obtained in the case of classical shift-invariant filtering. The purpose of such a system is to produce a high-level description of a 3-D scene on the base of its 2-D representation (or representations). In other words, we restrict our attention to the

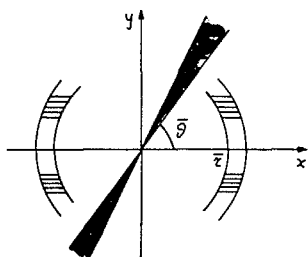

 ANALYSE DE SCENES PAR DES OPERATEURS FORME-INVARIANTS  
 SCENE ANALYSIS BY MEANS OF FORM-INVARIANT OPERATORS


Fig. 1

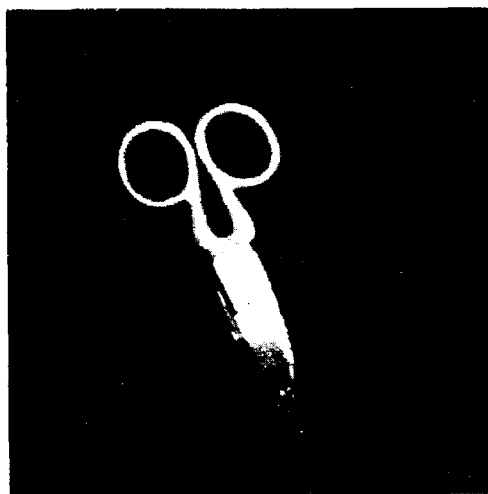
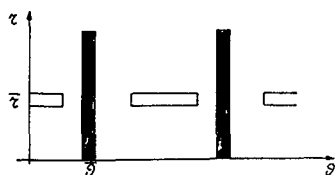


Fig. 2

Fig. 1 A possible multiprocessor algorithm implementing the cartesian to polar coordinates mapping via a raster scan access to the input and output data. Each processor captures and maps the input data falling in a small angular sector (dark area) or a small circular strip (white area).

Fig. 2 An original image obtained by digitizing the output of a TV-camera, and displayed (8 bit/pel) on the screen of the VDS-701 image processing system.

so called low vision stages, where mainly data-driven processings take place to yield the building blocks needed at higher level stages.

Among these low level operations we consider: a) quality improvements of the acquired image (such as contour enhancement); b) recognition of image parts (e.g. specific objects) representing the components of the scene description. We make the following assumptions and simplifications. First of all, the acquisition of 3-D objects by means of 2-D patterns is affected by parameter changes concerning viewpoint, illumination, noise, scale, rotation and occlusions. We consider here only scale, rotation and illumination. The scale and rotation of the available 2-D pattern depend, of course, on the (unknown) distance and relative position (forgetting the perspective) of the objects and the observer. The environment illumination affects the luminance gradients of the different surfaces constituting the image. It is therefore apparent the need for a preprocessing step of the acquired image aimed to feed the subsequent recognition stage with an input more suitable to produce recognition results less sensitive to the aforementioned parameters. Since recognition techniques are available that are scale and rotation independent, the preprocessing step must be devised so that i) the overall processing (filtering and recognition) is still scale (and rotation) invariant, ii) the noise and illumination effects are, at least partly, compensated. The first condition can be met if and only if the preprocessing itself is scale (and rotation)

invariant, i.e. if it belongs to the class of operators described in section 2. To satisfy condition ii), suitable weighting functions can be chosen: for noise smoothing purposes such functions will generally be of low-pass type, whereas for contour enhancement purposes (so that the effects of luminance changes can be reduced) they will have high-pass characteristics.

As regards the scale-independent recognition methods, they can be roughly divided into two broad classes: iconic techniques based on template matching and syntactic procedures. We assume that a template matching algorithm is used, in a suitable computational space where scale and rotation changes have the only effect to shift the correlation peak: the amounts of the shifts along the coordinate axes are a measure of the degree of relative scaling and rotation of the pattern with respect to the template. To summarize, our approach implies: segmenting the image in different patterns (potential objects); finding a center point of each pattern (e.g. the center of mass); form-invariant preprocessing; template matching.

The first experimental result we present concerns the form-invariant filtering of the 2-D pattern shown in Fig. 2. Such an image is processed by a shift-variant operator belonging to the subclass defined in eq. (4) and exhibiting a band-pass characteristic very similar to the operators described in /6,9/. The spatial resolution of this filtering decreases with increasing eccentricity, i.e. going from the center to the periphery (in all the examples the center is the center of

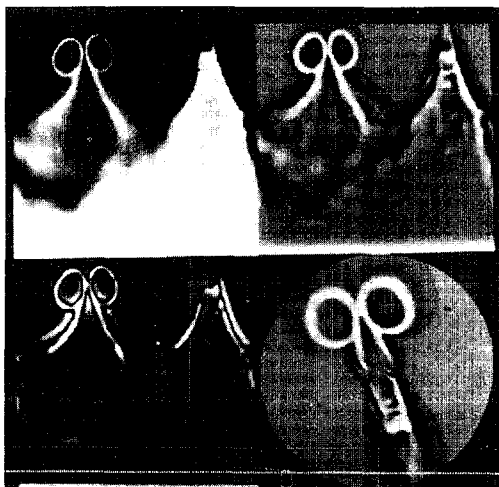


Fig. 3

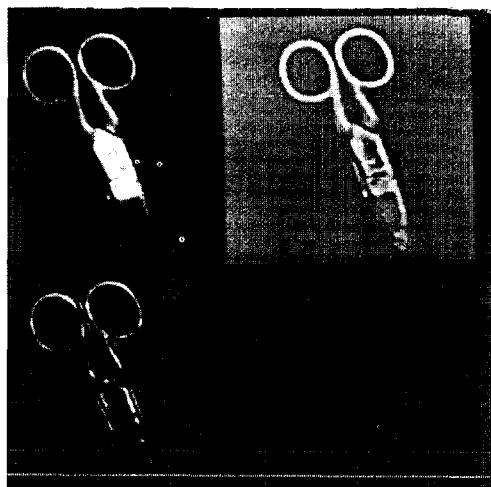


Fig. 4

Fig. 3 Result of a form-invariant processing (bottom right) performed by a set of circularly symmetrical band-pass filters having size linearly increasing from the center to the periphery. This result is also shown in the  $(\ln r, \vartheta)$  plane (top right) together with the original image (top left) and the output modulus (bottom left).

Fig. 4 Result (top right) of a shift-invariant processing obtained by means of an average size weighting function of the set used in Fig. 3, applied to the same original image (top left). The bottom left image is the output modulus.

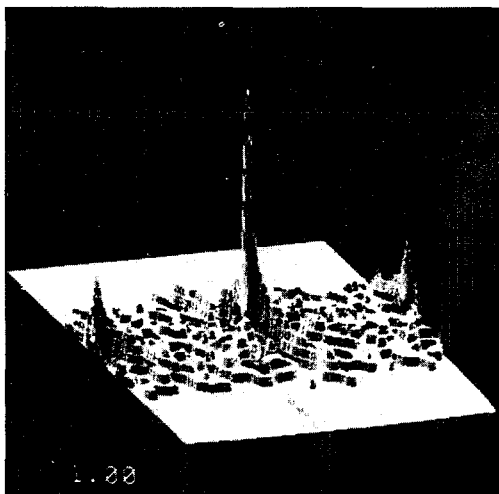


Fig. 5

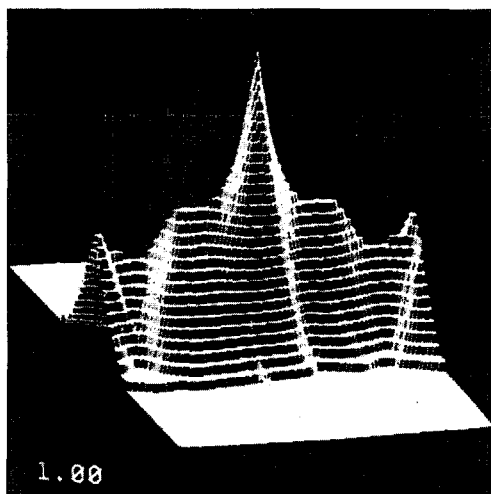


Fig. 6

Fig. 5 Autocorrelation of the scale-invariant contour-enhanced image of Fig.3, computed and displayed in the  $(\ln r, \vartheta)$  plane.

Fig. 6 Autocorrelation of the shift-invariant filtered image of Fig. 4, computed and displayed in the  $(\ln r, \vartheta)$  plane.

mass of the pattern). The result of this processing is shown in Fig. 3: the bottom right part of the figure shows that a non uniform contour reinforcement has taken place, being the contours more finely enhanced in the central area. The other parts of the figure show the original image (top left) and the output in the  $(\ln r, \vartheta)$  plane, using two different display techniques: a linear scale is used at top right (where the low-to-high intensity transitions are distinguishable from the high-to-low intensity transitions), while the

modulus is presented at bottom left. For comparison purposes, the result of a shift-invariant processing is presented in Fig. 4 (top right), that has been obtained by using the weighting function with average size in the set of form-invariant operators used to obtain Fig. 3. Of course, the contour enhancement is now uniform all over the image.

The difference between the processing results of Figs. 3 and 4 can be better appreciated if we consider the output images as potential templates in a scale-independent



C.Braccini, G.Gambardella, A.Grattarola  
ANALYSE DE SCENES PAR DES OPERATEURS FORME-INVARIANTS  
SCENE ANALYSIS BY MEANS OF FORM-INVARIANT OPERATORS

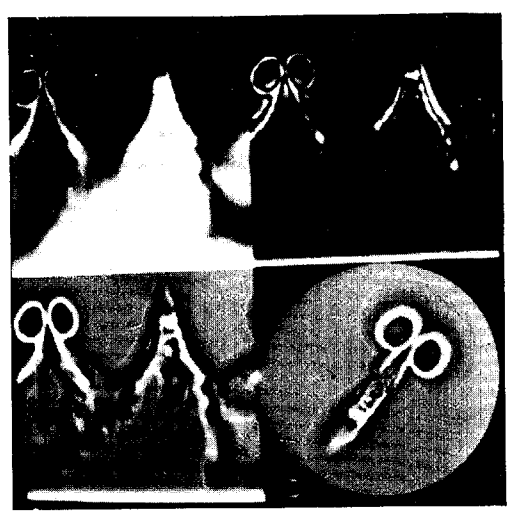


Fig. 7

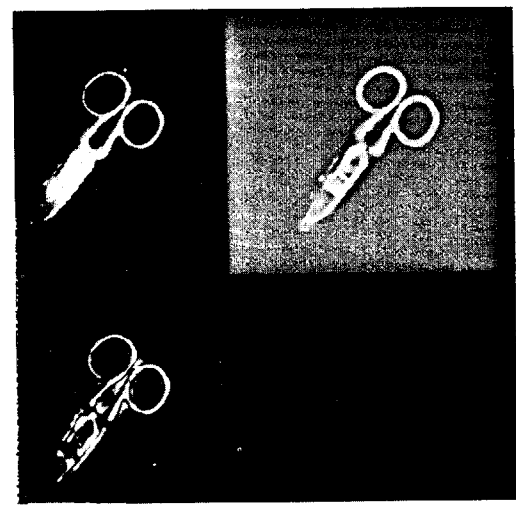


Fig. 8

Fig. 7 Result of the same form-invariant processing performed in Fig. 3 applied to a scaled and rotated version of the same original image. The output is shown in various forms, as in Fig. 3, except for the interchange of the top right and the bottom left quadrants.

Fig. 8 Shift-invariant filtering of the scaled and rotated version of the original image, shown as in Fig. 4.

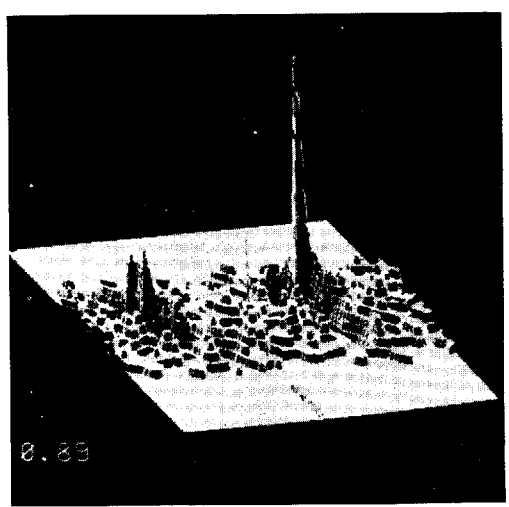


Fig. 9

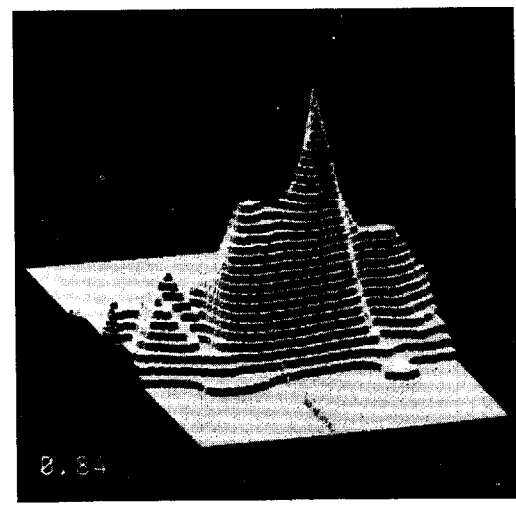


Fig. 10

Fig. 9 Cross-correlation of the output images of Figs.s 3 and 7, performed in the  $(\ln r, \theta)$  plane. The peak, whose normalized value is 0.89, occurs at a point in the  $(\ln r, \theta)$  plane corresponding to the rotation and scaling factors.

Fig. 10 Cross-correlation of the output images of Figs.s 4 and 8, performed in the  $(\ln r, \theta)$  plane. The peak position is the same as in Fig. 9 and its normalized value is 0.84.

template matching technique and compute their autocorrelation sequences in the  $(\ln r, \theta)$  plane. These latter are shown in Figs.s 5 and 6, respectively: the higher energy concentration near the origin in the autocorrelation of the scale-invariant output (Fig. 5) implies less noise sensitivity in the classification process. Such a different behavior is due to the fact that the details in Fig. 3 are more finely enhanced around the center of mass in the x-y plane, i.e. in an area

that covers a large portion of the  $(\ln r, \theta)$  plane and therefore heavily affects the autocorrelation sequence.

Similar results are obtained if the patterns to be matched differ in scale and rotation. Figures 7 and 8 show the outputs of the form-invariant and space-invariant processings, respectively, applied to a scaled and rotated version of the image of Fig. 2. We expect that the matching of the (filtered) scaled and unscaled patterns will give better



(i.e. more discriminating) results when the outputs of the form-invariant operators are correlated. This is easily checked by comparing the correlation sequence of the contour-enhanced images of Fig.s 3 and 7, shown in Fig. 9, with the correlation of the images of Fig.s 4 and 8 (Fig. 10). Again, the first correlation exhibits a more "peaky" behavior, with a larger peak value with respect to Fig. 10. In both cases, the peak position is a measure of the relative rotation and scale factors.

#### 4. CONCLUSIONS

The examples shown confirm that form-invariant operators produce patterns more suitable to scale-invariant recognition (based on the template matching techniques) than shift-invariant filters. This statement deserves further investigation, to cover a wider class of filters and to include the analysis of noise sensitivity.

#### ACKNOWLEDGEMENTS

This work has been supported by the National Research Council (C.N.R.) of Italy.

#### REFERENCES

- /1/ H.G. Barrow, J.M. Tenenbaum, "Computation al vision", Proc. of the IEEE, vol. 69, n. 5, May 1981, pp. 572-595.
- /2/ P.S. Schenker, "Toward the robot eye: isomorphic representation for machine vision", Proc. of SPIE, vol. 283, 3-D Machine Perception, 1981, pp. 30-47.
- /3/ C. Braccini, G. Gambardella, "Linear shift-variant filtering for form-invariant processing of linearly scaled signals", Signal Processing, vol. 4, n. 2-3, April 1982, pp. 209-213.
- /4/ J.G. Verly, "Circular and extended circular harmonic transforms and their relevance to image reconstruction from line integrals", J. Opt. Soc. Am., vol. 71, n. 7, July 1981, pp. 825-835.
- /5/ G. Robbins, T. Huang, "Inverse filtering for linear shift-variant imaging systems", Proc. of the IEEE, vol. 60, n. 7, July 1972, pp. 862-872.
- /6/ C. Braccini, G. Gambardella, G. Sandini, "A signal theory approach to the space and frequency variant filtering performed by the human visual system", Signal Processing, vol. 3, n. 3, July 1981, pp. 231-240.
- /7/ C.F.R. Weiman, G. Chaikin, "Logarithmic spiral grids for image processing and display", Comp. Graph. and Image Proc., vol. 11, n. 3, Nov. 1979, pp. 197-226.
- /8/ D. Casasent, D. Psaltis, "Deformation invariant, space-variant optical pattern recognition", in Progress in Optics, E. Wolf Ed., vol. XVI, North-Holland, 1978, pp. 291-356.
- /9/ C. Braccini, G. Gambardella, A. Grattarola, "The form-invariant 2-D filtering and its applications to pattern recognition", Proc. of EUSIPCO-83, Erlangen, September 12-16, 1983, pp. 207-210.