ADAPTIVE RIESZ BASIS DECOMPOSITION FOR IMAGE SEARCH

*Licia Capodiferro**, Marco Carli*, Luca Costantini*, Alessandro Neri*, Veronica Palma**

*Applied Electronics Dept.,Università degli Studi Roma TRE, Roma, Italy
**Fondazione Ugo Bordoni, Roma, Italy
email: lcapodiferro@fub.it, carli@uniroma3.it, lcostantini@fub.it, neri@uniroma3.it, vpalma1@uniroma3.it
web: www.comlab.uniroma3.it, www.fub.it

ABSTRACT

A novel technique via Quadtree Gauss-Laguerre Transform (LGT) for a complete localization of a complex image in an multimedia database is presented. This technique is based on an iterative Maximum Likelihood procedure that allows to compare a region of interest (ROI) of an image with a content of a database, independently from location, rotation and scale. At this aim the ROI is expanded in terms of a Riesz basis, consisting of a set Circular Harmonic Gauss Laguerre functions, selected in order to make a trade off between accuracy and computational complexity. Numerical results obtained by a Monte Carlo simulation illustrate the performance of the method.

1. INTRODUCTION

The accurate localization of patterns is an important issue in many image and video processing applications. For instance, estimation of position, rotation and scale of a given template within a complex scene, is a fundamental task in multimedia database management systems (DBMS), in indoor and outdoor surveillance, in human interactions recognition and classsification, in automatic vehicle guidance and in robotic applications, just to mention a few. Conventional solutions are based on invariants, i.e. on object features that are unaffected by one or more localization parameters. Typically, the pattern position t is estimated using a correlation-based invariant matching procedure [1],[2], and subsequently rotation and scale are estimated with respect to the already known position. Using invariants, suboptimum solutions from the accuracy viewpoint are obtained, but the complexity of the problem reduces to search in parameter spaces of lower dimension.

In this paper, a novel method based on a Riesz hypercomplete basis whose elements are the Gauss-Laguerre Circular Harmonic functions (CHFs) is presented. Gauss-Laguerre CHFs are complex, polar separable filters characterized by harmonic angular shape, a useful property to build rotationally invariant descriptors. Local expansions based on Gauss-Laguerre CHFs have been already introduced for Maximum Likelihood orientation invariant pattern recognition, [3]. However, a rather large number of expansion terms has to be employed when dealing with large objects containing many details. Here, to reduce the computational complexity, we propose to partition a region of interest of an image into smaller and smaller square blocks whose content is approximated by a truncated expansion making use of just a few Gauss Laguerre CHFs. The block width is controlled by the norm of the approximation error. To further reduce the computational complexity of the Maximum Likelihood estimation implementation, the elements of the quadtree blocks are ranked with respect to their amount of Fisher's information on location and rotation, proportional to the energy of the low pass filtered gradient. Then, a procedure based on the sequential matching of each block of the ranked quadtree list is applied.

2. THE PROPOSED METHOD

2.1 Gauss-Laguerre Transform

Let $\mathbf{x} = [x_1, x_2]$ be the coordinates in the real plane $\Re^2$. Any image $f(\mathbf{x}) \in L^2(\Re^2, d^2\mathbf{x})$ admits an orthogonal expansion under a Gaussian weighting function, $w(\mathbf{x}) = e^{\pi|\mathbf{x}|^2}$, complete over the entire plane, around any point $\xi = (\xi_1, \xi_2)$, [3]:

$$f(\mathbf{x})\, w\left(\frac{\mathbf{x}-\xi}{s}\right) = \sum_{n=0}^{\infty}\sum_{k=0}^{\infty} D_{n,k}(\xi)\,\frac{1}{s}\mathscr{L}_k^{(n)}\left(\frac{|\mathbf{x}-\xi|}{s}, \theta(\mathbf{x}-\xi)\right)$$

where $\theta()$ denotes the angular coordinate defined by the relationship

$$\theta(\mathbf{x}) = tg^{-1}\left(\frac{x_2}{x_1}\right),$$

$\mathscr{L}_k^{(n)}(r,\theta)$ are the *Gauss-Laguerre* functions defined as:

$$\mathscr{L}_k^{(n)}(r,\theta) = (-1)^k\, 2^{(|n|+1)/2}\pi^{|n|/2}\left[\frac{k!}{(|n|+k)!}\right]^{1/2} \cdot r^{|n|} L_k^{(n)}\left(2\pi r^2\right) e^{-\pi r^2} e^{jn\theta},$$

$L_k^{(n)}$ are generalized Laguerre polynomials defined by Rodriguez's formula:

$$L_k^{(n)}(t) = \frac{t^{-n}e^t}{k!}\frac{d^k}{dt^k}\left[t^{k+n}e^{-t}\right] = \sum_{h=0}^{k}(-1)^h \binom{n+k}{k-h}\frac{t^h}{h!},$$

and the expansion coefficients are defined as:

$$D_{n,k}(\xi) = \left\langle f(\mathbf{x})w\left(\frac{\mathbf{x}-\xi}{s}\right), \frac{1}{s}\mathscr{L}_k^{(n)}\left(\frac{|\mathbf{x}-\xi|}{s}, \theta(\mathbf{x}-\xi)\right)\right\rangle.$$

We incidentally observe that the parameter s controls the width of the weighting function.

The expansion in terms of Gauss-Laguerre functions can be derived, for instance, by first applying the Fourier's series expansion to the representation of the image f in polar coordinates with respect to the angular coordinate, and then expanding the radial profile of each harmonic using the Laguerre polynomials $L_k^{(n)}(t)$.

The Gauss-Laguerre functions are members of the wider class of Circular Harmonic Functions (CHFs), successfully used for many low level vision tasks, thanks to their selectivity with respect to basic visual patterns, [4],[5]. CHFs of n-th order are, by definition, polar separable functions of the form $h(r)e^{jn\theta}$.

By virtue of their harmonic angular shape, CHFs are indeed natural detectors for different classes of features: CHFs of order n=1 for example are tuned to edges, n = 2 to lines, n = 3 to forks, etc..

In addition, every Gauss-Laguerre function generates a dyadic Circular Harmonic Wavelet. This means that every image $f(\mathbf{x})$ can be represented by its continuous wavelet transform $W_{\mathscr{L}_k^n}[f](\mathbf{b},\alpha,\sigma)$ where $\mathbf{b}$, α and σ are the parameters representing respectively the translated, rotated and scaled version of the mother wavelet $\mathscr{L}_k^{(n)}$.

With reference to localization of complicated patterns, a rather relevant property is the following.

PROPERTY I. *Given an image f defined over a finite support $I \subset \Re^2$ and a lattice $\Xi = \{\xi_m \in I, m = 1,...,M\}$ the set of Gauss-Laguerre functions*

$$\left\{\frac{1}{s}\mathscr{L}_k^{(n)}\left(\frac{|\mathbf{x}-\xi_m|}{s},\theta(\mathbf{x}-\xi_m)\right), m = 1,...,M\right\}$$

defines a Riesz basis for f.

Proof. The orthogonality of the *Gauss-Laguerre* functions implies that

$$\sum_m \int_I \left| w\left(\frac{\mathbf{x}-\xi_k}{s}\right)\right|^2 |f(\mathbf{x})|^2 d\mathbf{x} = \sum_m\sum_n\sum_k |D_{n,k}(\xi_m)|^2,$$

therefore

$$\gamma\|f(\mathbf{x})\|^2 \le \sum_m\sum_n\sum_k |D_{n,k}(\xi_m)|^2 \le \Gamma\|f(\mathbf{x})\|^2,$$

with

$$\gamma = \min_{\mathbf{x}} \sum_{m=1}^M \left| w\left(\frac{\mathbf{x}-\xi_k}{s}\right)\right|^2,$$

and

$$\Gamma = \sum_{m=1}^M \int_I \left| w\left(\frac{\mathbf{x}-\xi_k}{s}\right)\right|^2 d\mathbf{x}.$$

q.e.d.

Thus in turn implies that the inner product between two images f and g with expansion coefficients $D_{n,k}(\xi_m)$ and $C_{n,k}(\xi_m)$, respectively, satisfies the following condition:

$$\frac{\gamma}{\Gamma}\langle f(\mathbf{x}),g(\mathbf{x})\rangle \le \frac{1}{\Gamma}\sum_m\sum_n\sum_k D_{n,k}(\xi_m)C^*_{n,k}(\xi_m) \le \langle f(\mathbf{x}),g(\mathbf{x})\rangle$$

The magnitude of the approximation error strictly depends on the ratio γ/Γ, that can be a priori computed. Moreover, it could be demonstrated that the more general set of Gauss-Laguerre functions

$$\left\{\frac{1}{s_k}\mathscr{L}_k^{(n)}\left(\frac{|\mathbf{x}-\xi_m|}{s_k},\theta(\mathbf{x}-\xi_m)\right), m = 1,...,M\right\}$$

it is a Riesz basis too. Since, in order to reduce the computational complexity only a limited number of expansion coefficients are considered, here we propose the quadtree decomposition adaptive scheme for choosing the lattice Ξ and the shape parameters $\{s_k\}$ realizing a good trade off between accuracy and complexity.

Since the adaptive scheme is applied, here, for pattern recognition and localization, let us first express the Likelihood functional corresponding to the estimation of pattern translation, rotation and scale, in terms of the said Riesz basis.

2.2 Maximum Likelihood Localization

Let now $f(\mathbf{x})$ be the observed region of interest that contains a noisy, translated, rotated and scaled copy of a given template pattern $g(\mathbf{x})$ so that we have:

$$w[R_\varphi(\mathbf{x}-\mathbf{b})]f(\mathbf{x}) = w[R_\varphi(\mathbf{x}-\mathbf{b})]g\left[R_\varphi\left(\frac{\mathbf{x}-\mathbf{b}}{a}\right)\right] + v(\mathbf{x}),$$

where the parameters a, $\mathbf{b}$ and φ represent respectively scale, position and rotation of the observed image, $v(\mathbf{x})$ is the observation noise that can be modeled as a sample function drawn from a white, zero-mean Gaussian random field with power density spectrum equal to $N_0/4$ and R_φ is the rotation matrix defined as:

$$R_\varphi = \begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix}.$$

Let $\theta = [\mathbf{b},a,\varphi]$ be the unknown parameter vector, the Likelihood functional is given by the conditional probability of f w.r.t. θ, divided by any arbitrary function that does not depend on θ :

$$\ln\Lambda[f(\mathbf{x});\mathbf{b},a,\varphi] = -\frac{2}{N_0}\int\int\left|w\left[R_\varphi\left(\frac{\mathbf{x}-\mathbf{b}}{a}\right)\right]\right|^2 \times \\ \times\left|f(\mathbf{x}) - g\left[R_\varphi\left(\frac{\mathbf{x}-\mathbf{b}}{a}\right)\right]\right|^2 d\mathbf{x}.$$

Direct evaluation of the Maximum Likelihood solution presents a rather high computational complexity because the search of the maximum for $\mathbf{b}$, φ and a implies an exhaustive search in a four dimensional space. However the choice of LG functions as expansion basis and the choice of gaussian window which is rotation invariant leads to a simpler iterative procedure,[6] requiring an exhaustive search just in a two dimensional space. In fact, considering that any n-th order CHF can be steered in any direction φ by simple multiplication by the complex factor $e^{-jn\varphi}$, and denoting with $\eta_{n,k}(\mathbf{x};a)$ the expansion coefficients of $g(\mathbf{x}/a)$, we can approximate the ML functional as follows:

$$\ln\Lambda[f(\mathbf{x});\mathbf{b},a,\varphi] \simeq const + \\ -\frac{2}{N_0\Gamma}\sum_m\sum_n\sum_k\left|D_{n,k}(\xi_m) - \eta_{n,k}\left[R_\varphi\left(\frac{\xi_m-\mathbf{b}}{a}\right);a\right]e^{-jn\varphi}\right|^2.$$

On the other hand, denoting with $C_{n,k}(\mathbf{x})$ the expansion coefficients of $g(\mathbf{x})$ for $a = 1$ (i.e. $C_{n,k}(\mathbf{x}) = \eta_{n,k}(\mathbf{x};1)$) the following interscale relationship holds

$$\eta_{n,k}(\mathbf{x};a) = \sum_{l=k}^{\infty} B(a;n,k,l) C_{n,k}(\mathbf{x}),$$

where

$$B(a;n,k,l) = (-1)^{l-k} \sqrt{\frac{(n+l)!l!}{(n+k)!k!}} \cdot \frac{a^{-n-2k}}{(l-k)!} \left(1 - \frac{1}{a^2}\right)^{l-k}, (l \geq k).$$

Thus, the ML functional can be further approximated as follows:

$$\ln \Lambda[f(\mathbf{x}); \mathbf{b}, a, \varphi] \simeq const - \frac{2}{N_0 \Gamma} \times \sum_m \sum_n \sum_k \left| D_{n,k}(\xi_m) - \sum_{l=k}^{\infty} B(a;n,k,l) C_{n,k} \left[R_\varphi \left(\frac{\xi_m - \mathbf{b}}{a} \right) \right] e^{-jn\varphi} \right|^2$$

The maxima of the expression of the above functional w.r.t. scale a and orientation φ represent the *Gauss-Laguerre Likelihood Map* (GLLM):

$$GLLM(\mathbf{b}) = \max_{a,\varphi} \{ \ln \Lambda [f(\mathbf{x}); a, \varphi, \mathbf{b}] \}$$

The local estimate of the maxima can been performed by means of quasi-Newton maximization procedure as the Broyden-Fletcher-Goldfarb-Shanno algorithm. The location of the absolute maximum of this map provides the estimated position of the pattern. The resulting *Gauss-Laguerre Likelihood Map* indicates, point by point, the best matches between the two images under all possible orientations and scales.

2.3 Quadtree Decomposition

Since $f(\mathbf{x})$ may contain multiple objects with arbitrary shape, direct use of Gauss Laguerre expansion as well of other CHFs expansions, as those in Zernike's moments,for computing the ML functional would require a larger and larger number of expansion terms. Thus,in order to reduce the computational complexity, we resort to the hypercomplete Riesz basis that allows to partition the region of interest into smaller squares, so that for each of them a truncated Gauss-Laguerre expansion with a reduced number of terms can be utilized. Using a *quadtree* decomposition, a block is further subdivided if the norm of the approximation error between the reconstructed image and the image itself exceeds a predefined threshold. In fact, the smaller is the support, the smaller is the approximation error when a finite number of terms is employed.

More in detail, let R represent the region of interest, eventually coincident with the whole image, and let P be a predicate equal to *True* whenever the accuracy of the approximation of the current Riesz basis can be considered satisfactory. R is partitioned into smaller and smaller square regions $R^{(i)}$, so that for each $R^{(i)}$, $P(R^{(i)}) = True$. Initially the basis set is empty and the current region $R^{(0)}$ is set equal to the given ROI. At the i-th step of the recursion, the center ξ_i of the current region $R^{(i)}$ is evaluated and the subset of functions

$$\left\{ \frac{1}{s_i} \mathscr{L}_k^{(n)} \left(\frac{|\mathbf{x} - \xi_\mathbf{i}|}{s_i}, \theta(\mathbf{x} - \xi_\mathbf{i}) \right), k = 1, ..., K, n = 1, ..., N \right\}$$

is added to the current basis set as a potential candidate. Then the predicate P is computed. If $P(R) = False$ the current candidate subset is removed from the basis set, the current region is split into four squares and the quadtree decomposition is recursively applied to each of them. If P is *True* the current candidate subset is definitively added to the basis and the current region is not split anymore.

In order to control the computational complexity of the whole procedure, we chose as predicate P the comparison of the L^2 norm of the approximation error in the reconstruction of a square block of the image with a predefined number of Gauss Laguerre coefficients with a threshold t. If the norm of the error between the image itself $f(\mathbf{x})$ and the reconstructed image $\hat{f}(\mathbf{x})$ using the current basis exceeds a predefined threshold, P is set to false. Let us denote with δ_i the width of $R^{(i)}$, and with $w_T(\mathbf{x})$ a square window of unitary width, then

$$P(R^{(i)}) = \left\{ \left\| w_T \left(\frac{\mathbf{x} - \xi_i}{\delta_i} \right) [f(\mathbf{x}) - \hat{f}(\mathbf{x})] \right\|^2 < t \right\}.$$

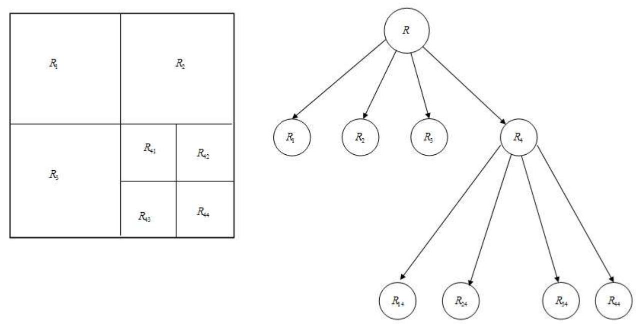


Figure 1: Quadtree structure

Pattern location, rotation and scale estimation accuracy is strictly related to the Fisher's information. However, as demonstrated in [3] this quantity is proportional to the magnitude of the energy of the derivatives along two orthogonal directions and to the energy of the angular derivative, or, equivalently, to the effective spatial and angular bandwidths.

Therefore, in order to design a sequential detection and estimation procedure that verifies whether each candidate image contains each square of the quadtree, in order to reduce the search time, we first rank the template quadtree blocks on the basis of the energy of the mid and high, angular and radial, frequency components, computed directly from the Gauss Laguerre expansion coefficients.

As an alternative, salient points based on invariants can be extracted and quadtree blocks can then be ranked based on the saliency of the key points falling inside them, [7].

When a ROI of a given image has to be searched in a database, the first block of the ranked list is considered and each database candidate image is expanded in Gauss-Laguerre coefficients using the same base employed for the current quadtree block.

Since rotation of a pattern simply produces a linear phase shift of each expansion coefficient proportional to the order of the angular harmonic, detection of the pattern belonging to the first square of the ranked quadtree list can be performed by means of a quasi-Newton maximization procedure as the Broyden-Fletcher-Goldfarb-Shanno algorithm maximizing, for each **b** the quantity

$$GLLF^{(1)}(\mathbf{b},a,\varphi) = -\frac{2}{N_0\Gamma}\times \sum_{n=0}^{N}\sum_{k=0}^{K}\left|D_{n,k}(\xi_c)-\sum_{l=0}^{L}B(a;n,k,l)C_{n,k}(\xi_c-\mathbf{b})e^{-jn\varphi}\right|^2$$

where ξ_c denotes the center of the current region.

Thus, for each discrete location of a grid, the rotation and the scale maximizes the $GLLF^{(1)}$ functional are determined and then a discrete direct search is performed to determine its absolute maximum. Thus, at the first step the parameter estimate is

$$[\hat{\mathbf{b}}^{(1)},\hat{a}^{(1)},\hat{\varphi}^{(1)}] = Arg\left\{\max_{\mathbf{b},a,\varphi}\left[GLLF^{(1)}(\mathbf{b},a,\varphi)\right]\right\}$$

Once for each image of the dataset the local maximum of $GLLF^{(1)}$ has been computed, the images are ranked on the basis of this absolute maximum. Then the image corresponding to the highest $GLLF^{(1)}$ is selected as the potential candidate for image matching, and $[\hat{\mathbf{b}}^{(1)},\hat{a}^{(1)},\hat{\varphi}^{(1)}]$ is employed as coarse estimate in order to verify whether the candidate image contains the second block of the rank ordered list of quadtree elements, too.

With respect to the first block, the $GLLF^{(2)}$ map is built only for a limited set of possible locations, falling inside a small neighbor of the site predicted on the basis of the coarse estimates. In addition, the quasi-Newton procedure utilized to maximize $GLLF^{(2)}$ is initialized using the coarse estimate too.

If the energy of the difference between the subset of the reference template, constituted by the first and the second square of the quadtree and the current image falls below a predefine threshold, location and rotation of the image are refined and the next square analyzed.

In general, at the h-th stage the $GLLF^{(h)}$ map is computed using the first h points of the lattice, ranked according to the saliency indicator, i.e.,

$$GLLF^{(h)}[\mathbf{b},a,\varphi] = -\frac{2}{N_0\Gamma}\times\sum_{m}\sum_{n}\sum_{k} \left|D_{n,k}(\xi_m)-\sum_{l=k}^{\infty}B(a;n,k,l)C_{n,k}\left[R_\varphi\left(\frac{\xi_m-\mathbf{b}}{a}\right)\right]e^{-jn\varphi}\right|^2$$

The procedure ends when the last block in the list has been processed. If at some stage the energy of the difference exceeds a predefined threshold, the current image is discarded and the next item of the dataset corresponding to the highest $GLLF^{(1)}$ is considered as candidate for pattern matching.

3. EXPERIMENTAL RESULTS

The proposed method has been tested on a 52 images database. For each grey-level image, we have 4 different views (256x256 pixels) corresponding to different orientations and scales, the first one is the original image, the second image is scaled, the third one is rotated and the last one is scaled and rotated. See fig.2. In the performed simula-

Figure 2: Three database samples with different orientations

tions, the Gauss-Laguerre expansion has been truncated to the (angular) order n=6, and to the (radial) order k=7. This gives, for each quadtree block a descriptor array of 173 elements. The value s_m of the weighting gaussian window is matched to quadtree block size. In fig.(3) an example of the Likelihood map for the "Einstein" image is showed. In table (1) some results on angle and scale estimate error for some images from the multimedia database are showed. The angle and scale estimate errors are quite low and the algorithm is capable to find the searched points in the candidate image, estimating rotation and scale of the image with a low error rate.

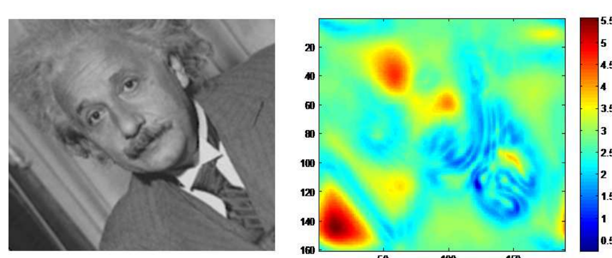


Figure 3: Gauss-Laguerre Likelihood map of "Einstein" image

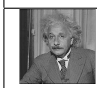
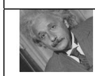
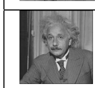
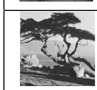
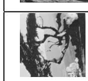
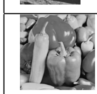
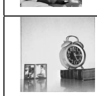
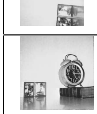
Image	Angle Estimate Error [deg]	Scale Estimate Error
 airplane	0	0
 airplane-rot10-scal90	0.19	0.01
 airplane-rot15	0.58	0.01
 airplane-scal92	0.07	0.02
 einstein	0	0
 einstein-rot40-scal97	2.58	0.03
 einstein-rot220	3.13	0
 einstein-scal97	0	0.03
 tree	0	0
 tree-rot195-scal87	0.38	0.01
 tree-rot35	0.09	0.01
 tree-scal80	0.11	0.01
 peppers	0	0
 peppers-rot78-scal95	5.81	0.07
 peppers-rot98	1.65	0.01
 peppers-scal82	0.78	0.07
 clock	0	0
 clock-rot90-scal92	0.05	0.01
 clock-rot85	0.22	0.08
 clock-scal98	0.59	0.01

Table 1: Angle and scale estimate error for some images from the database

4. CONCLUSIONS

In this paper a novel technique for template matching based on Laguerre-Guass Transform and on a Quadtree decomposition has been presented. The Gauss-Laguerre Transform allows a simpler iterative Likelihood functional estimate compared to the traditional Maximum Likelihood based on searching the image with the whole set of rotated and scaled images. In particular, it allows an approximated ML solution with not an expensive computational cost. Thanks to the Gaussian windowing, this method is well suited for localization of patterns of complex objects. The experimental results show an high detection rate and an accurate location estimate and show how this class of Circular Harmonic filters performs very well in presence of scaling and rotation.

REFERENCES

[1] Young Deok Chun, Nam Chul Kim, Ick Hoon Jang, “Content-Based Image Retrieval Using Multiresolution Color and Texture Features,” in *IEEE Transactions on-Multimedia*, Vol. 10, Issue 6, Page(s):1073 - 1084, Oct. 2008.

[2] Fei Li; Qionghai Dai; Wenli Xu; G. Er , “Multilabel Neighborhood Propagation for Region-Based Image Retrieval,” in *IEEE Transactions on Multimedia*, Vol. 10, Issue 8, Page(s):1592 - 1604 ,Dec. 2008

[3] A.Neri and G. Iacovitti, ”Maximum Likelihood Localization of 2-D Patterns in the Gauss-Laguerre Transform Domain: Theoretic Framework and Preliminary Results,”,*Image Processing, IEEE Transactions on*,vol. 13, pp.72-86, Jan. 2004.

[4] L. Sorgi, N. Cimminiello, and A. Neri, “Keypoint selection in the laguerre-gauss transformed domain,” in *Proc. of 2nd Workshop on Applications of Computer Vision*, ECCV, May 2006.

[5] Z. Zalevsky and I. Ouzieli, and D. Mendlovic,”Wavelet-transform-based composite filters for invariant pattern recognition”, *Applied Optics*, vol.35, pp. 3141-3147, Jun. 1996.

[6] M. Carli, F. Coppola, G. Iacovitti, and A. Neri, “Translation, orientation and scale estimation based on laguerre-gauss circular harmonic pyramids,” *SPIE Conf.Photonics West* , 2002.

[7] L. Capodiferro and E. D. Di Claudio and G. Jacovitti and F. Mangiatordi,”Application of Local Fisher Information Analysis to Salient Points Extraction” in*Proceedings of the IASTED International Conference on Signal Processing, Pattern Recognition, and Applications*, SPPRA 2008, 13-15 February, Innsbruck, Austria