

Non photorealistic rendering in frequency domain

F.Mangiatoridi, E.Pallotti, V.Baroncini, L.Capodiferro
Fondazione Ugo Bordoni, Viale del Policlinico 147, 00161 Roma, Italy

Abstract

In this work the authors present a novel image abstraction and stylization framework based on the analysis of natural scene in the multi-resolution Laguerre Gauss (LG) domain. The extraction of complex LG image sketches at different resolutions and the corresponding non linear adjustment ensure the preservation of the shapes and the deleting of the unimportant edge details. The reduction of the color range within the region interiors is obtained applying the image LG synthesis formula to the smoothed simplified image sketches. The proposed method is computationally easy to implement and provides the possibility to perform a fine tune regulation of the stylization parameters, customizing the NPR process. Subjective assessment of the proposed method were designed to compare it with two low level "anchors"; the complete set of subjective tests are in progress; preliminary expert viewing test show some effectiveness of the proposed NPR technique.

Introduction

Over the last decades the computer graphics and image/video processing communities have produced an extensive research in the field of non-photorealistic rendering (NPR) for images and video to create stylized and abstract depictions which convey some aspects of natural scenes more effectively [1–4]. The main goal is to provide visual communication designer and illustrator with computer aided tools to support their work and retain human creativity [5, 6]. In fact automatic and semi-automatic image stylization algorithms can be applied in a variety of applications ranging from visual communication to digital multimedia technology such as advertising, comics, video games, image and video animations, augmented reality.

To address this challenge, several sophisticated methodologies are proposed in literature. An exhaustive survey on the state of the NPR art is presented by Kyprianidis with a specific focus on the recent methods for transforming 2D contents (photographs and videos) into artistic stylizations [1, 7].

Generally non-photorealistic rendering involves two different phases : line extraction and region smoothing [8]. The first aims to capture the significant height discontinuities by which the image objects are sketched, the second reduces the color range and removes insignificant details. Following Winnemoller approach, many NPR framework implement these two task with separate filtering operations in spatial domain [9].

In this paper the authors propose a simple but effective image abstraction process in the multi-resolution frequency domain of Laguerre Gauss (LG) wavelet. To generate the cartoon-like image, low frequency component and high frequency components are extracted from the photo and manipulated at dyadic resolutions before being manipulated to generate the cartoon-like image. In fact it is shown [10] that the LG plane of first angular order locally represents the edge lines of the boundary regions defining the

image shapes. However the hyper-completeness of LG image representation ensures the perfect image reconstruction from the multiscale smoothed complex gradients. This yields the authors to develop an NPR system from the manipulation of LG image coefficients.

This contribution is arranged as follow. The second section reviews the relevant properties of LG multi-resolution image decomposition and outlines the theoretical results from which the present work was originated. The third section describes the LG-wavelet based NPR system, focusing on the main technical details of the all modules involved. The fourth section describes the two evaluation methodologies adopted to evaluate the attractiveness of the cartoonized images comparing the proposed method with two low level "anchors"; some preliminary experimental results are also provided. Finally, conclusions and future directions are given in the fifth section.

Theoretical Background

The conceptual bases of the proposed NPR-method is the Laguerre-Gauss (LG) wavelet transform which provides a general purpose framework for the rotation invariant detection of image patterns and structures.

The LG wavelets are complex,circular harmonic orthonormal functions, that can be written in polar coordinates (r, θ) as:

$$g_{n,k}(r, \theta; s) = h_n^{(k)}(r; s) \cdot e^{jn\theta} \quad (1)$$

with the radial profile $h_n(r; s)$ defined as:

$$h_n(r; s) = \frac{(-1)^k}{\sqrt{k!(n+k)!}} \left(\frac{r}{s}\right)^n \frac{1}{s\sqrt{2\pi}} L_k^n \left[\left(\frac{r}{s}\right)^2\right] e^{-\frac{1}{2}\left(\frac{r}{s}\right)^2} \quad (2)$$

where $L_k^n(x)$ is the generalized Laguerre polynomial defined by Rodriguez's formula:

$$L_k^n(x) = \sum_{i=0}^k (-1)^i \binom{n+k}{k-i} \frac{x^i}{i!} \quad (3)$$

An interesting property of the LG wavelet functions is that they are steerable and radially isomorphous with their Fourier spectrum, given by

$$\tilde{G}_{n,k}(\rho, \gamma) = \frac{(-j)^n}{\sqrt{k!(n+k)!}} (s\rho)^n L_k^n [(s\rho)^2] e^{-\frac{1}{2}(s\rho)^2} \cdot e^{jn\gamma} \quad (4)$$

with (ρ, η) polar coordinates in Fourier domain.

Laguerre-Gauss CHF's constitutes an orthonormal complete system [11], so any image $I(x,y)$, under a Gaussian weighting function, can be locally represented as:

$$I_G(r, \theta; s) = \sum_n \sum_k A_{n,k} g_n^{(k)}(r, \theta; s) \quad (5)$$

LG based NPR framework

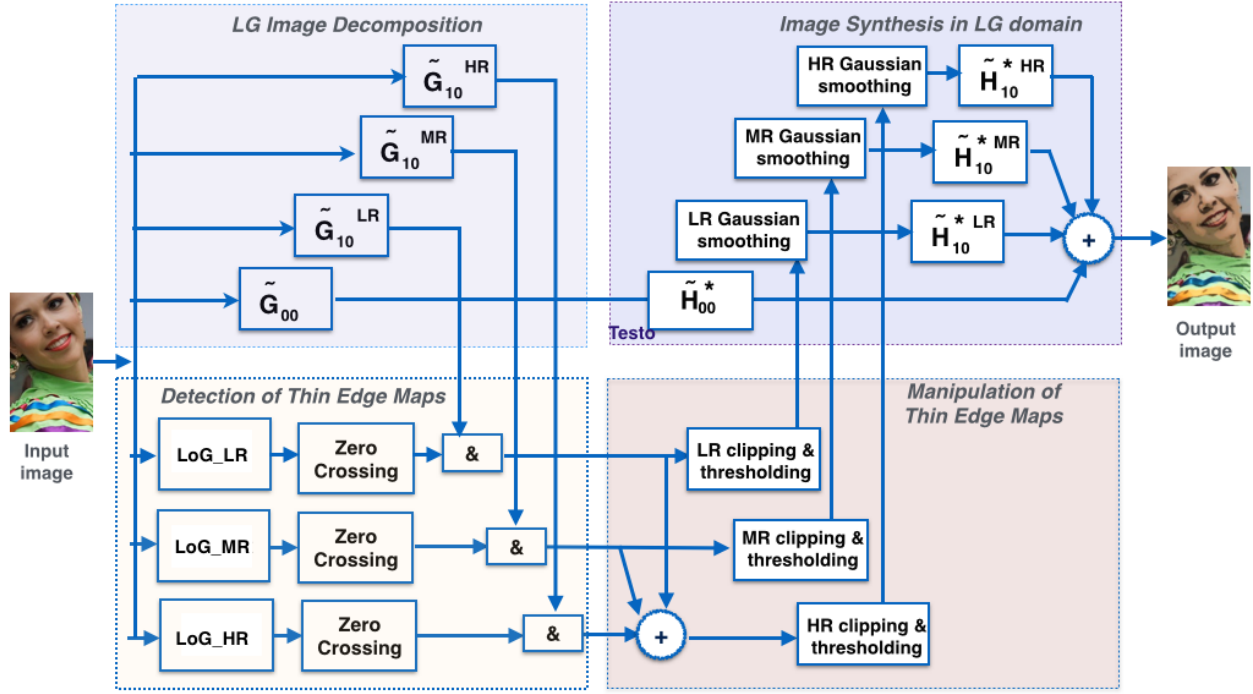


Figure 1: Framework for the proposed LG non-photorealistic rendering method

The coefficients A_{nk} in (5) represent the LG wavelet transform of the image and can be computed by means of the 2-D convolution.

$$A_{n,k} = I(r, \theta) * g_n^{(k)}(r, \theta; s) \quad (6)$$

This means that the LG coefficients A_{nk} can also be extracted filtering the image with the bank of LG steerable filters.

$$\tilde{A}_{n,k}(\xi, \eta; s) = \tilde{G}_{n,k}(\xi, \eta; s) \cdot \tilde{I}(\xi, \eta) \quad (7)$$

where $\tilde{I}(\xi, \eta)$ and $\tilde{A}_{n,k}(\xi, \eta; s)$ are the Fourier transform of $I(r, \theta)$ and $A_{n,k}$. In the decomposition formula of eq. (5), the parameter s defines the neighborhood where the image structural information is extracted.

Given a discrete dyadic sequences of scale $s_i = \sigma_0 2^i$, the corresponding set of LG functions $\{g_{n,k}\}$ define a hypercomplete dictionary of waveforms that allows the multiresolution representation of the image linearly combining the dictionary elements [11]. In particular, the original image can be perfectly synthesized from LG planes $A_{n,k}$ for $n = 0, 1$ and $k = 0$ at three different dyadic scales $\{s_i = \sigma_0 2^i, i = 0, 1, 2\}$ by the following reconstruction expression in the frequency domain:

$$\begin{aligned} \tilde{I} &= \frac{\sum_{i=0}^2 \tilde{A}_{1,0}(\xi, \eta; s_i) \cdot \tilde{G}_{10}^*(\xi, \eta; s_i) + \tilde{A}_{0,0}(\xi, \eta; s_0) \cdot \tilde{G}_{00}(\xi, \eta)}{\sum_{i=0}^2 |\tilde{G}_{10}^{(0)}(\xi, \eta; s_i)|^2 + |\tilde{G}_{00}^{(0)}(\xi, \eta; s_0)|^2} \\ &= \sum_{i=0}^2 \tilde{A}_{1,0}(\xi, \eta; s_i) \cdot \tilde{H}_{10}(\xi, \eta; s_i) + \tilde{A}_{0,0}(\xi, \eta; s_0) \cdot \tilde{H}_{00}(\xi, \eta) \end{aligned} \quad (8)$$

LG based NPR framework

The proposed approach to non-photorealistic rendering involves four main steps:

- *LG image decomposition* for the extraction of the image local structural information at different resolutions;
- *detection of the thin edge maps* for the fine characterization of the relevant contour lines and boundary regions within the image;
- *manipulation of the edge sketches* to increase the visual distinctiveness of the main shapes while deleting the irrelevant details and texture;
- *image synthesis* to generate the simplified cartoon-like images by interpolating the manipulated edge sketches in LG domain

The proposed NPR rendering framework is visualized in Fig. (1), which is further detailed in the following subsections.

LG image decomposition

The central problem of NPR task is the reduction of scene complexity while preserving its important structures. In the proposed technique, this issue is addressed in LG domain for the interesting property of LG wavelet transform to be intrinsically tuned to the various image patterns (edges, lines, crosses, texture, etc.) [12].

In particular, to fully characterize the structural information with the minimal set of elements, it is assumed to decompose the image in four LG components: a low frequency coefficient $A_{0,0}$ and three high frequency coefficients of first angular order $\{A_{1,0}(\xi, \eta; s_i)\}$ with $s_i = \sigma_0 2^i$ for $\{i = 0, 1, 2\}$.

The high frequency LG components $A_{1,0}(\xi, \eta; s_i)$ constitute

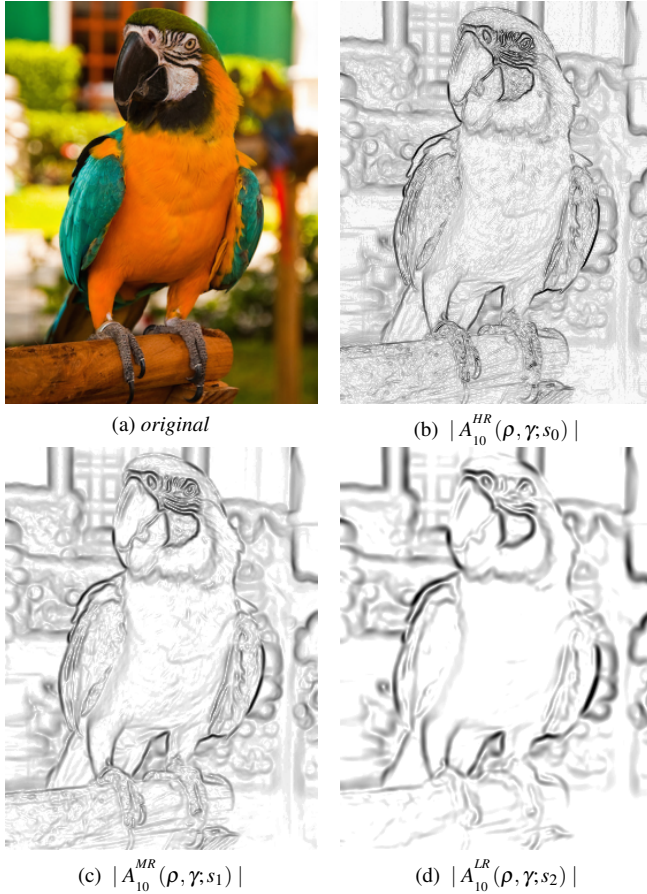


Figure 2: Original image (a) and the modulus of its complex gradient at high resolution (b), medium resolution (c), low resolution (d)

the image complex gradients at three different resolution scales, denoted respectively as high resolution (HR), medium resolution (MR) and low resolution (LR) scale.

These LG coefficients are extracted by processing the original photo with a bank of four LG CH steerable filters defined by eq. (4) for $s = s_i, k = 0$ and $n = 0, 1$, as shown in fig. (1).

Detection of thin edge maps

The first order LG coefficients A_{10} capture height discontinuities of the image $I(\rho, \gamma)$. When a large scale s is employed, the coefficient A_{10} detects the sharp signal variations associated with the contour of the objects within the image. On the contrary, fine edge details, associated with small amplitude fluctuations, can be extracted for small values of the scale s .

This implies that the coefficients $A_{10}(\rho, \gamma; s_i)$ are useful to extract the simplified image sketches conveying the meaningful information at each resolution. In fact, the image edge are defined as the local maxima of the modulus of the complex gradient A_{10} in the direction of $\angle A_{10}(\rho, \gamma; s)$.

To reach the goal of image abstraction, a simplified gradient wavelet representation is considered by sampling the LG coefficients $A_{10}(\rho, \gamma; s_i)$ in correspondence of the edge points for each different resolution scale. This operation generates the Complex Multi-resolution Sketch (CMS), defined as the set of the

sampled complex LG gradients of the image. The modulus of the CMS components constitutes the multi-resolution thin edge maps, that conveys only the relevant image structures.

In this work, the exact position of the edges is located following the Marr-Hildreth approach [13]. Therefore, the original image is convolved with the Laplacian of Gaussian operator $LoG(\sigma_{LP}^i)$ for each resolution scale to obtain the corresponding binary map of Laplacian zero crossing $BM_{\sigma_{LP}^i}$.

These binary maps are put in AND with the first order wavelet coefficients $A_{10}(\rho, \gamma; s_i)$, to generate the complex sketch $CS_{\sigma_{LP}^i}(\rho, \gamma; s_i)$ and the corresponding thin edge maps $|CS_{\sigma_{LP}^i}(\rho, \gamma; s_i)|$.

The quantity of the meaningful information gathered by the thin edge maps at high, medium, low resolution scale is strictly correlated with the sensitivity parameter of laplacian operator σ_{LP} and the wavelet scale σ_0 .

In the proposed NPR system, this parameter is set to default values for each scale resolution. However the user is given the capability to modulate σ_{LP} at each scale resolution to control the quantity of edge details in the CMS maps.



Figure 3: HR thin edge map $|A_{10}(\rho, \gamma; s_i)|$ and the modulus of its HR complex sketch for three decreasing values of laplacian sensitivity $\sigma_{LP}^0, \sigma_{LP}^1, \sigma_{LP}^2$

Manipulation of the thin edge maps

The clearness and abstraction of shapes within the image is operated by applying clipping and stretching operation on the

complex sketches at high, medium and low resolution. Each pixel value z of the complex sketch at medium and low resolution is transformed non linearly by scaling its modulus with $f(z)$ while preserving its orientation at each resolution s_i .

$$f(z) = \begin{cases} 0 & \text{for } |z| \leq \alpha \\ G \cdot (|z| - \alpha) & \text{for } \alpha \leq |z| \leq \beta \\ \max\{A_{10}(\rho, \gamma; s_i)\} & \text{otherwise} \end{cases} \quad (9)$$

where $\beta = \max\{A_{10}(\rho, \gamma; s_i)\} / G$.

To highlight the main region contours within the cartoonized image, the HR complex sketch is summed to MR and LR complex sketches; subsequently the modulus of obtained map is clipped and stretched with $f(z)$ operator. The gain G and the threshold α are selected by the user for each resolution scale (HR, MR, LR); the α parameter is set as a percentile value, taking into account the modulus of the complex sketches. The tuning of G and α parameters results in different stylization effects on the reconstructed image.

Image synthesis in LG domain

The cartoon like image is generated using the LG reconstruction formula (8) where the original coefficients A_{10} are replaced with synthetic versions Y_{10} calculated from the manipulated complex sketch \tilde{CS}_{10} .

$$\tilde{I}_{cartoon}(\xi, \eta) = \sum_{i=0}^2 \tilde{Y}_{1,0}(\xi, \eta; s_i) \cdot \tilde{H}_{10}^*(\xi, \eta; s_i) \quad (10)$$

For each resolution the coefficient $Y_{10}(\rho, \gamma; s_i)$ is obtained smoothing the manipulated complex sketch with a gaussian filter $G_s(\xi, \eta; \sigma_c) = \beta \cdot G_{0,0}(\xi, \eta; \sigma_c)$

$$\tilde{Y}_{1,0}(\xi, \eta; s_i) = \tilde{G}_s(\xi, \eta; \sigma_c) \cdot \tilde{CS}(\xi, \eta; s_i) \quad (11)$$

In fact, referring to an image with a vertical unit step edge, it is shown that its LG coefficient $A_{1,0}$ can be perfectly calculated from the corresponding complex sketch \tilde{CS} in frequency domain as:

$$\tilde{A}_{1,0}(\xi, \eta; s_i) = \tilde{G}_0^{(0)}(\xi, \eta; s_i) \cdot \tilde{CS}(\xi, \eta; s_i) \quad (12)$$

A large variety of NPR effects can be created by properly adjusting the parameters β and σ_c at each scale resolution.

Experimental results

The LG-based NPR system was preliminary tested on both monochrome and color images with different values of the parameters (the kernel-size of LoG, the low and upper limits of the stretching operator, the size of gaussian smoothing filter) that control the simplification process of the edge maps. The visual comparison of the obtained results respect to the original photos has shown the effectiveness of the proposed technique. In fact the LG-based NPR method gains a good level of abstraction without losing the image content comprehension and its pleasantness. Figures (4-6) show some relevant results for the proposed NPR technique. Other examples are available on line [14].

However, it should be noted that the exhaustive complete evaluation of experimental results is one of the most difficult



(a) Original image



(b) LG-NPR output

Figure 4: Non photo realistic rendering in LG domain for "Balu dog" image

problem in the NPR area [15], and some different methodologies are proposed in literature [16–19].

In the design of our experiments we have referred to the Wallraven approach for the evaluation of the subjective appeal of the abstracted images [18] asking users about their impression on the abstracted imagery. For this reason we designed a test in which the participation of expert and naive users evaluate the attractiveness of the LG-based non-photorealistic rendering results.

The test was designed including, as low level anchor, two sets of stylized images generated by means of commercial tools (i.e. ToyViewer 5.4.0, PhotoShop CS4). The images used in the test are selected from a dataset (test-DB) of 60 photographs, collected using the holiday and tourists tags on Flickr. These photo selection criteria ensured the due diversity in the dataset, in terms of image contents and photographer styles.

All the selected photos were cartoonized using the LG-NPR method and ToyViewer and PhotoShop tools, applying for each



(a) Original image



(b) LG-NPR output

Figure 5: Non photo realistic rendering in LG domain for "Vietnamese Boy" image



(a) Original image



(b) LG-NPR output

Figure 6: Non photo realistic rendering in LG domain for "London Bridge" image

tool three different set of calibration parameters. This led, for each image a set of 9 stylized pictures.

The first visual assessment experiment was designed to involve fifteen users, 8 males and 6 females aging from 18 to 35, half experts in the graphical art area, and half naive.

The assessment is made of two separate sessions where the same cartoonized pictures are presented to the users in pseudo-random, in a way to obtain two different evaluation of each test point. A pair of images is displayed side by side on the centre of the screen in a mid grey background. The left image is always the original while the right image is the cartoonized sample.

The scoring of the images is done using an 11 grades quality scale ranging from 0 to 10, where 0 is the lowest level of appreciation and 10 is the highest appreciation degree.

The Mean Opinion Score (MOS) and the Standard Deviation are computed to rank the three tools. Table 1 shows the obtained results.

In the second visual assessment experiment we visually compared the cartoonized versions of each test photo. In particular

we selected the cartoonized images that for each tool were rated with the highest scores in the first experiment.

An example of the three abstract versions of the same photo is shown in Fig.7(b), Fig.8 (a) and Fig.8(b).

The selected images are further evaluated using a no reference subjective assessment test to foresee the best cartoonized version. The same fifteen viewers, participating to the first visual assessment experiment, are involved.

For each input photo, we presented on the screen the three abstract images in random order and the users did not know which tool is adopted to generate the versions in specific positions. For each

Table 1: Mean and standard deviation of the normalized score obtained by each NPR tool

Tool	Average	Std dev.
ToyViewer	5.672	2.056
PS CS4	7.118	1.542
LG-NPR system	8.125	1.368



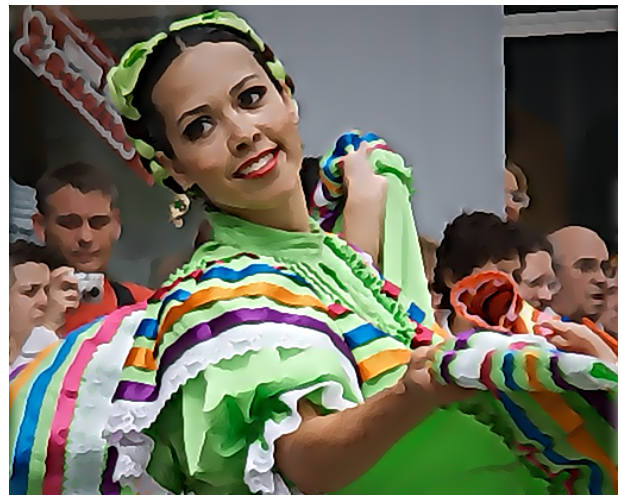
(a) Original image



(a) PS version



(b) ToyViewer version



(b) LG-NPR version

Figure 7: Original (a) and ToyViewer NPR version (b) of "Spanish Girl" image

Figure 8: PS cartoonized version (a) and LG-NPR version (b) of "Spanish Girl" image

group of three images the user selected the most attractive. A score (1) is assigned to the tool used to process the selected image while the rest receive zero vote. The ranking of the tools is then made with a simple counts of the assigned scores. Table 2 shows the obtained results.

Table 2: Percentual Score

Tool	Average
ToyViewer	16.08%
PS CS4	38.68%
LG-NPR system	45.24%

Conclusion

This paper proposes a novel technique for non-photorealistic rendering based on Laguerre Gauss local representation of the image. The effectiveness of the proposed approach was demonstrated by conducting visual comparison of the obtained result with the ones of some commercial tools. A systematic user evaluation of the attractiveness of the proposed method is also performed focusing the attention on a limited set of calibration parameters to obtain the same experimental conditions. The LG- NPR technique can be also used to realize a video cartoonization task after converting the video to individual image frames. However the current implementation of LG- based NPR system doesn't support the real time processing but a parallel multicore implementation may be required.

References

- [1] J.E. Kyprianidis, J.E., J. Collomosse, Tinghuai Wang, Isenberg, T., "State of the 'Art': A Taxonomy of Artistic Stylization Techniques for Images and Video," *Visualization and Computer Graphics*, IEEE Transactions on , vol.19, no.5, pp.866,885, May 2013
- [2] Henry Kang, Seungyong Lee, Charles K. Chui, "Flow-Based Image Abstraction", *IEEE Transactions on Visualization & Computer Graphics*, vol.15, no. 1, pp. 62-76, January/February 2009.
- [3] H. Winnemöller, S. C. Olsen, and B. Gooch, "Real-time video abstraction", in *ACM SIGGRAPH 2006 Papers, SIGGRAPH 06*, (New York, NY, USA), pp. 1221-1226, ACM, 2006.
- [4] D. DeCarlo and A. Santella, "Stylization and abstraction of photographs", in *Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 02*, (New York, NY, USA), pp. 769-776, ACM, 2002.
- [5] C. Reynolds, Stylized Depiction in Computer Graphics, <http://www.red3d.com/cwr/npr/>, 2012.
- [6] M. Zhao and S.-C. Zhu, Sisley the Abstract Painter, *Proc. Intl Symp. Nonphotorealistic Animation and Rendering (NPAR)*, pp. 99-107, 2010.
- [7] J. Collomosse and J.E. Kyprianidis, Artistic Stylization of Images and Video, Tutorial at Eurographics, 2011.
- [8] H.Kang, S.Lee, and C.Chui, "Flow-based image abstraction", *Visualization and Computer Graphics*, vol.15, no. 1, pp.62-76, 2009.
- [9] H. Winnemöller, J. E. Kyprianidis, and S. C. Olsen, Xdog: An extended difference-of-gaussians compendium including advanced image stylization, *Computers & Graphics*, vol. 36, pp. 740-753, 10 2012.
- [10] Capodiferro, L.; Di Claudio, E.D.; Iacolucci, F.; Laurenti, A.; Jacovitti, G., "Two-channel technique for high dynamic range image visualization," *Information Visualisation*, 2004. IV 2004. Proceedings. Eighth International Conference on , vol., no., pp.269,273, 14-16 July 2004
- [11] G. Jacovitti, A. Neri, "Multiresolution circular harmonic decomposition", *IEEE Trans. on Signal Processing*, Vol.. 48. No. 11, pages 3242-3247, November 2000.
- [12] A. Neri, G. Jacovitti, Maximum Likelihood Localization of 2D Patterns in the Gauss-Laguerre Transform Domain: Theoretic Framework and Preliminary Results, *IEEE Trans. on Image Processing*, Vol. 13, Issue: 1, Pages. 72-86, Jan. 2004.
- [13] Marr, D., Hildreth, E., Theory of edge detection, *Proc. Roy. Soc. London*, vol. 207, No. 1167, pp. 187-217.
- [14] <http://www.yesresearch.eu/LG-NPR.html>
- [15] A. Hertzmann, "Non-photorealistic rendering and the science of art," in *International Symposium on Non-Photorealistic Animation and Rendering*, Annecy: New York: ACM, 2010, pp. 147-157.
- [16] A. Santella, The art of seeing: Visual perception in design and evaluation of non-photorealistic rendering, Ph.D. dissertation, Rutgers University, 2005. [Online]. Available: <http://www.research.rutgers.edu/asantell/thesis.pdf>
- [17] H. Winnemöller, Perceptually-motivated non-photorealistic graphics, Ph.D. dissertation, Northwestern University - Illinois, 2006. [Online]. Available: http://holgerweb.net/PhD/Research/papers/diss_winnemoeller06.pdf
- [18] C. Wallraven, H. H. Bulthoff, D. W. Cunningham, J. Fischer, and D. Bartz, "Evaluation of real-world and computer-generated stylized facial expressions", *ACM Transactions on Applied Perception*, vol. 4, no. 3, pp. 16:1-16:24, 2007.
- [19] De Alencar, J.B.O.; De Queiroz, J.E.R.; Gomes, H.M., "An Approach for Non-photorealistic Rendering That Is Appealing to Human

Viewers," *Graphics, Patterns and Images (SIBGRAPI)*, 2013 26th SIBGRAPI - Conference on , vol., no., pp.242,249, 5-8 Aug. 2013

Author Biography

Vittorio Baroncini is bacca Laureate in Physics on digital filters from the University of Rome "La Sapienza", Italy. From 1986 he works at Fondazione Bordini dealing with objective and subjective quality assessment and services evaluation. He has a long experience in the International Standardization area (ITU and ISO), also chairing several technical groups. His research interest regards acceptance metrics and services evaluation protocols for new Digital TV and CINEMA standards (3D and HDR).

Licia Capodiferro received her Dr. Ing. degree in Electronic Engineering from the University of Rome "La Sapienza", Italy. In 1987 she joined the Fondazione Ugo Bordini where she currently works as head of the Department of Information Processing and Management. Her main research interests are in the field of multimedia processing, with a focus on algorithms that allow the use of images and videos on the different types of terminals.

Federica Mangiatordi received the M.Sc. Degree in Electronic Engineering at University of Rome La Sapienza and the PhD in Electronic Materials, Optoelectronics and Microsystems from the University of Roma TRE. She works at Fondazione Ugo Bordini from 2007. Her research interest concern multimedia retrieval, image restoration algorithms, novel metrics for full reference and no-reference image objective quality assessment.

Emiliano Pallotti received the "Laurea" Degree in Telecommunications Engineering at the University of Rome "La Sapienza", Italy, and PhD in Electronic Materials, Optoelectronics and Microsystems from the University of Roma TRE.

In 2007 he joined the Fondazione Ugo Bordini where his research activities are in the field of on computational algorithms and video processing techniques based on multiresolution image representation in wavelet domain.