The Gaussian transform

Alecu, Teodor; Voloshynovskyy, Svyatoslav; Pun, Thierry

THE GAUSSIAN TRANSFORM

Teodor Iulian Alecu, Sviatoslav Voloshynovskiy and Thierry Pun

Computer Vision and Multimedia Laboratory, University of Geneva, 24 Rue Général-Dufour, 1204 Geneva, Switzerland .
phone: + (41) 22 379 1084, fax: + (41) 22 379 7780, email: Teodor.Alecu@cui.unige.ch

ABSTRACT

This paper introduces the general purpose Gaussian Transform, which aims at representing a generic symmetric distribution as an infinite mixture of Gaussian distributions. We start by the mathematical formulation of the problem and continue with the investigation of the conditions of existence of such a transform. Our analysis leads to the derivation of analytical and numerical tools for the computation of the Gaussian Transform, mainly based on the Laplace and Fourier transforms, as well as of the afferent properties set (e.g. the transform of sums of independent variables).
Finally, the Gaussian Transform is exemplified in analytical form for typical distributions (e.g. Gaussian, Laplacian), and in numerical form for the Generalized Gaussian and Generalized Cauchy distributions families.

1. INTRODUCTION

Gaussian distributions are extensively used in the (broad sense) signal processing community, mainly for computational benefits. For instance, in an estimation problem Gaussian priors yield quadratic functionals and linear solutions. In rate-distortion and coding theories, closed form results are mostly available for Gaussian source and channel descriptions. Real data, however, is generally not Gaussian distributed. The goal of the work presented in this paper is to describe non-Gaussian distributions through an infinite mixture of Gaussian distributions.
In a related work [4], it was proven that any distribution can be approximated through a mixture of Gaussian up to an arbitrary level of precision. However, no hint was given by the author on how to obtain the desired mixture in the general case. In [5], an analytical formula is given for an infinite mixture of Gaussians equivalent to the Laplacian distribution, and used in a source coding application. Unfortunately, no generalization was attempted by the authors. The work presented here has the roots in their proof and extends the concept to a wide range of symmetric distributions through the introduced Gaussian Transform.
The Gaussian Transform concept and the results presented in this paper can be extensively used in various applications of signal and image processing and communications including estimation, detection, source and channel coding etc.
The rest of the paper is divided in two main blocks. In Section 2 we define the Gaussian Transform, analyze its existence, investigate its properties and derive the mathematical tools for analytical and/or numerical computation, whereas in Section 3 we exemplify both the transform for some typical distributions such as Generalized Gaussian and Generalized Cauchy, and some of the properties deduced in Section 2.

2. GAUSSIAN TRANSFORM

2.1 Definition and existence

We consider a generic symmetric distribution *p(x)*. As we are aiming at representing it through an infinite mixture of Gaussians, we can safely disregard the mean, and assume for simplicity reasons that *p(x)* is zero-mean. We are looking for an integral representation in the form:

$$\int_0^\infty G\left(\sigma^2\right)\mathcal{N}\left(x\mid\sigma^2\right)d\sigma^2 = p\left(x\right), \qquad (1)$$

where $\mathcal{N}\left(x\mid\sigma^2\right)$ is the zero-mean Gaussian distribution:

$$\mathcal{N}\left(x\mid\sigma^2\right)=\frac{1}{\sqrt{2\pi\sigma^2}}\mathrm{e}^{-x^2/2\sigma^2},$$

and $G\left(\sigma^2\right)$ is the mixing function that should reproduce the original *p(x)*. We can now introduce the Gaussian Transform.

Definition 1 *(Gaussian Transform)*. The direct Gaussian Transform $\mathcal{G}$ is defined as the operator which transforms $p\left(x\right)$ into $G\left(\sigma^2\right)$, and the Inverse Gaussian Transform $\mathcal{G}^{-1}$ is defined as the operator which maps $G\left(\sigma^2\right)$ to $p\left(x\right)$:

$$\mathcal{G}: p\left(x\right)\mapsto G\left(\sigma^2\right);\ \mathcal{G}^{-1}: G\left(\sigma^2\right)\mapsto p\left(x\right).$$

Obviously, $\mathcal{G}^{-1}$ is simply given by (1):

$$\mathcal{G}^{-1}\left(G\left(\sigma^2\right)\right)=\int_0^\infty G\left(\sigma^2\right)\mathcal{N}\left(x\mid\sigma^2\right)d\sigma^2. \qquad (2)$$

We look now for the direct Gaussian Transform. We need to prove that for a given $p\left(x\right)$ a mixing distribution $G\left(\sigma^2\right)$ exists such as to comply with (1). This can be summarized in three conditions.
Condition 1. For a given *p(x)*, a function $G\left(\sigma^2\right)$ defined according to (1) exists.
Condition 2. This function is non-negative.
Condition 3. Its integral $\int_0^\infty G\left(\sigma^2\right)d\sigma^2$ is equal to 1.
The last condition is a consequence of Condition 1. Indeed, if $G\left(\sigma^2\right)$ exists, then, integrating both sides of (1) with respect to *x* and inverting the integration order of the left side:

$$\int_{-\infty}^{\infty}\left(\int_0^{\infty} G(\sigma^2)\mathcal{N}(x\mid\sigma^2)d\sigma^2\right)dx = \int_{-\infty}^{\infty} p(x)dx, \text{ thus}$$
$$\int_0^{\infty}\left(G(\sigma^2)\int_{-\infty}^{\infty}\mathcal{N}(x\mid\sigma^2)dx\right)d\sigma^2 = 1.$$
Finally, since $\mathcal{N}(x\mid\sigma^2)$ is a distribution:
$$\int_0^{\infty} G(\sigma^2)d\sigma^2 = 1 .$$
In order to investigate Condition 1, the existence of $G(\sigma^2)$, perform the following variable substitutions: $s = x^2$ and $t = \frac{1}{2\sigma^2}$. Since *p(x)* is symmetric, it can be rewritten as:
$$p(x) = p(|x|) = p(\sqrt{s}) .$$
The left hand side of (1) transforms to:
$$\int_0^{\infty} G(\sigma^2)\mathcal{N}(x\mid\sigma^2)d\sigma^2 = \int_0^{\infty} G\left(\frac{1}{2t}\right)\frac{1}{2t}\frac{1}{\sqrt{\pi t}}e^{-st}dt .$$
According to the definition of the Laplace Transform $\mathcal{L}$ [1], equation (1) finally takes the form:
$$\mathcal{L}\left(G\left(\frac{1}{2t}\right)\frac{1}{2t}\frac{1}{\sqrt{\pi t}}\right) = p(\sqrt{s}). \quad (3)$$
Thus, $G(\sigma^2)$ is linked to the original probability distribution *p(x)* through the Laplace Transform and can be computed using the Inverse Laplace Transform $\mathcal{L}^{-1}$. The direct Gaussian Transform is therefore given by:
$$\mathcal{G}(p(x)) = \frac{1}{\sigma^2}\sqrt{\frac{\pi}{2\sigma^2}}\left(\mathcal{L}^{-1}\left(p(\sqrt{s})\right)(t)\right)_{t=\frac{1}{2\sigma^2}} . \quad (4)$$
Consequently, the existence of the Gaussian Transform is conditioned by the existence of the Inverse Laplace Transform of $p(\sqrt{s})$. From the general properties of the Laplace Transform, it is sufficient [1] to prove that the limit at infinity of $s\cdot p(\sqrt{s})$ is bounded, or equivalently :
$$\lim_{x\to\infty} x^2\cdot p(x) < \infty . \quad (5)$$
The above condition is satisfied by all the distributions from the exponential family, as well as by all the distributions with finite variance. (4) allows for straightforward identification of Gaussian Transforms for distributions whose Laplace Transforms are known, by simply using handbook tables.

Unfortunately, it does not guarantee compliance with the Condition 2: non-negativity. As it is rather difficult to verify a priori this constraint, the test should be performed a posteriori, either analytically or numerically.

2.2 Properties of the Gaussian Transform

We derive the first property of the Gaussian Transform using the initial value theorem for the Laplace Transform [1], the direct formula (4) and the existence condition (5).

Final Value Property. The Gaussian Transform tends asymptotically to 0 when σ^2 tends to infinity:
$$\lim_{\sigma^2\to\infty} G(\sigma^2) = 0 . \quad (6)$$
We also expect the transform to preserve the data variance.

Mean Value Property. The mean value of the Gaussian Transform is equal to the variance of the original distribution:
$$\int_{-\infty}^{\infty} x^2 p(x)dx = \int_0^{\infty}\sigma^2 G(\sigma^2)d\sigma^2 . \quad (7)$$
Proof: $\int_{-\infty}^{\infty} x^2 p(x)dx = \int_{-\infty}^{\infty} x^2\left(\int_0^{\infty} G(\sigma^2)\mathcal{N}(x\mid\sigma^2)d\sigma^2\right)dx =$

$= \int_0^{\infty} G(\sigma^2)\left(\int_{-\infty}^{\infty} x^2\mathcal{N}(x\mid\sigma^2)dx\right)d\sigma^2 = \int_0^{\infty} G(\sigma^2)\sigma^2 d\sigma^2$ q.e.d.

Another crucial property of the Gaussian Transform applies to the transform of the sum of independent variables.

Convolution Property. If X_1 and X_2 are independent random variables with $\mathcal{G}(p_{X_1}(x_1)) = G_{X_1}$ and $\mathcal{G}(p_{X_2}(x_2)) = G_{X_2}$, then the Gaussian Transform of their sum is the convolution of their respective Gaussian Transforms (the result can be generalized for the sum of multiple independent variables):
$$\mathcal{G}(p_{X_1+X_2}(x_1+x_2)) = G_{X_1} * G_{X_2} . \quad (8)$$
Proof: consider the random variable $X = X_1 + X_2$. Since G_{X_1} exists, X_1 is a random Gaussian variable with variance $\sigma_{X_1}^2$ distributed according to the distribution probability G_{X_1}. Similarly, X_2 is a random Gaussian variable with variance $\sigma_{X_2}^2$ distributed according to G_{X_2}. Then *X* is also a random Gaussian variable with variance $\sigma_X^2 = \sigma_{X_1}^2 + \sigma_{X_2}^2$. But $\sigma_{X_1}^2$ and $\sigma_{X_2}^2$ are independent variables drawn from G_{X_1} and G_{X_2}. It follows that σ_X^2 is a random variable described by the probability distribution $G_X = G_{X_1} * G_{X_2}$, q.e.d.

Corollary. If X_1 and X_2 are independent random variables and X_2 is Gaussian distributed $p_{X_2}(x_2) = \mathcal{N}(x_2\mid\sigma_{X_2}^2)$, then the Gaussian transform of their sum is a shifted version of G_{X_1}:
$$G_{X_1+X_2}(\sigma^2) = \begin{cases} 0, \sigma^2 < \sigma_{X_2}^2 \\ G_{X_1}(\sigma^2 - \sigma_{X_2}^2), \sigma^2 \geq \sigma_{X_2}^2 \end{cases} . \quad (9)$$

2.3 Numerical computation

The computation of the Gaussian Transform for distributions not available in handbooks is still possible through the complex inversion method for Laplace Transforms known as the Bromwich integral [3]:
$$\mathcal{L}^{-1}(f(s)) = \frac{1}{2\pi i}\int_{\varepsilon-i\infty}^{\varepsilon+i\infty} f(u)e^{ut}du , \quad (10)$$
where ε is a real positive constant satisfying $\varepsilon > \sup(\mathrm{Re}(poles(f)))$ and *u* is an auxiliary variable. (10) can be rearranged as a Fourier Transform, allowing the use of the numerous numerical and/or symbolical packages available (with ω the variable in the Fourier space):

$$\mathcal{L}^{-1}\left(f\left(s\right)\right)=e^{\varepsilon t}\mathcal{F}^{-1}\left(f\left(\varepsilon+i\omega\right)\right) \tag{11}$$

Very often *p(x)* has no poles, being a continuous and bounded function, and in this case it might be very practical to evaluate (11) at the limit $\varepsilon \to 0$. Using (4) and (11):

$$G\left(\sigma^2\right)=\frac{1}{\sigma^2}\sqrt{\frac{\pi}{2\sigma^2}}\left(\mathcal{F}^{-1}\left(p\left(\sqrt{i\omega}\right)\right)\right)_{t=\frac{1}{2\sigma^2}}. \tag{12}$$

When the original distribution is only numerically known, approximation of $G\left(\sigma^2\right)$ is still possible either through analytical approximations of *p(x)*, followed by (4) (or (12)), or through solving the inverse problem yielded by (1).

3. EXAMPLES OF GAUSSIAN TRANSFORMS

3.1 Analytic Gaussian transforms

The most obvious and natural example is the Gaussian Transform of a *Gaussian* distribution. One would expect to have (with δ the Dirac function):

$$\mathcal{G}\left(\mathcal{N}\left(x\mid\sigma_0^2\right)\right)=\delta\left(\sigma^2-\sigma_0^2\right). \tag{13}$$

Indeed, from (4):

$$\mathcal{G}\left(\mathcal{N}\left(x\mid\sigma_0^2\right)\right)=\left(t\sqrt{\frac{2t}{\sigma_0^2}}L^{-1}\left(e^{-s/2\sigma_0^2}\right)\right)_{t=\frac{1}{2\sigma^2}}.$$

Then, using the Laplace Transform tables:

$$\mathcal{G}\left(\mathcal{N}\left(x\mid\sigma_0^2\right)\right)=\frac{1}{2\sigma^2}\sqrt{\frac{1}{\sigma^2\sigma_0^2}}\delta\left(\left(t-\frac{1}{2\sigma_0^2}\right)\right)_{t=\frac{1}{2\sigma^2}}.$$

Since the Gaussian Transform is a function of σ^2 [2]:

$$\delta\left(\frac{1}{2\sigma^2}-\frac{1}{2\sigma_0^2}\right)=2\left(\sigma_0^2\right)^2\delta\left(\sigma^2-\sigma_0^2\right), \text{ and}$$

$$\mathcal{G}\left(\mathcal{N}\left(x\mid\sigma_0^2\right)\right)=\frac{\sigma_0^2}{\sigma^2}\sqrt{\frac{\sigma_0^2}{\sigma^2}}\delta\left(\sigma^2-\sigma_0^2\right)=\delta\left(\sigma^2-\sigma_0^2\right).$$

Thus the Gaussian Transform of a Gaussian distribution $\mathcal{N}\left(x\mid\sigma_0^2\right)$ is a Dirac function centered in σ_0^2.

The convolution property (8) can now be used to prove a well known result in statistics: the sum of two independent Gaussian variables, with respective probability laws $\mathcal{N}\left(x\mid\sigma_{X_1}^2\right)$ and $\mathcal{N}\left(x\mid\sigma_{X_2}^2\right)$, is another Gaussian variable with probability distribution $\mathcal{N}\left(x\mid\sigma_{X_1}^2+\sigma_{X_2}^2\right)$ (the extension to non-zero mean distributions is trivial). Proof:

$$\mathcal{G}\left(p_{X_1}\right)=\delta\left(\sigma^2-\sigma_{X_1}^2\right);\ \mathcal{G}\left(p_{X_2}\right)=\delta\left(\sigma^2-\sigma_{X_2}^2\right) \text{ thus}$$
$$\mathcal{G}\left(p_{X_1+X_2}\right)=\mathcal{G}\left(p_{X_1}\right)*\mathcal{G}\left(p_{X_2}\right)=\delta\left(\sigma^2-\left(\sigma_{X_1}^2+\sigma_{X_2}^2\right)\right).$$

Then, inverting the Gaussian Transform:

$$p_{X_1+X_2}=\mathcal{G}^{-1}\left(\delta\left(\sigma^2-\left(\sigma_{X_1}^2+\sigma_{X_2}^2\right)\right)\right)=\mathcal{N}\left(x\mid\sigma_{X_1}^2+\sigma_{X_2}^2\right).$$

Similarly to (13), it is possible to compute the Gaussian transforms of other usual symmetric distributions using the Laplace transform tables [1]. We exemplify with the *Laplacian* and *Cauchy* distributions.

Laplacian:

$$p_X^L\left(x\mid\lambda\right)=\frac{\lambda}{2}e^{-\lambda x};\ p_X^L\left(\sqrt{s}\mid\lambda\right)=\frac{\lambda}{2}e^{-\lambda\sqrt{s}}$$
$$\mathcal{G}\left(p_X^L\left(\sqrt{s}\mid\lambda\right)\right)=\frac{\lambda^2}{2}e^{-\frac{\lambda^2}{2}\sigma^2}. \tag{14}$$

As mentioned, the result (14) was already proven in [5].

Cauchy:

$$p_X^C\left(x\mid b\right)=\frac{1}{\pi}\frac{b}{b^2+x^2};\ p_X^C\left(\sqrt{s}\mid b\right)=\frac{1}{\pi}\frac{b}{b^2+s}$$
$$\mathcal{G}\left(p_X^C\left(x\mid b\right)\right)=\frac{1}{\sigma^2}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{b}{2\sigma^2}}. \tag{15}$$

The results (13), (14) and (15) are plotted in Fig. 2.

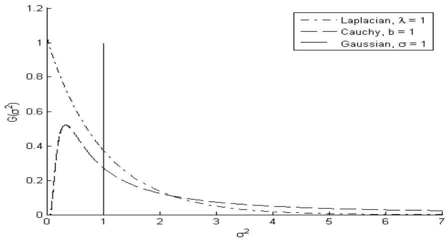

Figure 2. Gaussian Transforms

To exemplify the *Convolution Property*, consider now Cauchy data contaminated with independent additive white Gaussian noise. Then the Gaussian Transform of the measured data (Fig. 3) is a shifted version of the original data transform (9).

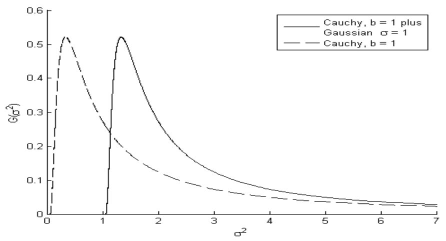

Figure 3. Shift of the Gaussian Transform

The previous analytical results will be generalized in subsection *B* through numerical computations.

3.2 Numeric Gaussian transforms

This part illustrates the numerical computation of the Gaussian Transform through (12) for the Generalized Gaussian and Generalized Cauchy distributions families.

Generalized Gaussian Distribution (GGD)

The GGD family is described by an exponential probability density function with parameters γ and σ_γ:

$$p_X^G\left(x\mid\gamma,\sigma_\gamma\right)=\frac{\gamma\eta\left(\gamma\right)}{2\Gamma\left(1/\gamma\right)}\frac{1}{\sigma_\gamma}e\left(-\left(\eta\left(\gamma\right)\left|\frac{x}{\sigma_\gamma}\right|\right)^\gamma\right),$$

where $\eta\left(\gamma\right)=\sqrt{\Gamma\left(3/\gamma\right)\Gamma\left(1/\gamma\right)^{-1}}$. For $\gamma=1$ the GGD particularizes to the Laplacian distribution, while for $\gamma=2$ one obtains the Gaussian distribution. The Gaussian Transform of the Generalized Gaussian distribution can not be obtained in analytical form using (4). However, it does exist (5) and can be calculated numerically through (12):

$$G_{X|\gamma,\sigma_\gamma}\left(\sigma^2\right)=\frac{1}{\sigma^2}\sqrt{\frac{\pi}{2\sigma^2}}\left(\mathcal{F}^{-1}\left(p^G_{X|\gamma,\sigma_\gamma}\left(\sqrt{i\omega}\right)\right)\right)_{t=\frac{1}{2\sigma^2}}$$

$$p^G_{X|\gamma,\sigma_\gamma}\left(\sqrt{i\omega}\right)=\frac{\gamma\eta(\gamma)}{2\Gamma\left(1/\gamma\right)}\frac{1}{\sigma_\gamma}e\left(-\left(\frac{\eta(\gamma)}{\sigma_\gamma}\right)^\gamma(i\omega)^{\gamma/2}\right).$$

The Gaussian Transforms for γ ranging from 0.5 to 2 with fixed $\sigma_\gamma=1$ are plotted in Fig. 6. The transforms evolve from a Dirac-like distribution centered on 0 for small γ to exponential for $\gamma=1$, then Rayleigh-like for $\gamma=1.2$, bell-shaped for $\gamma=1.5$ and again Dirac-like centered on σ_γ^2 for $\gamma=2$. As expected, the Gaussian Transform of the Laplacian distribution ($\gamma=1$) is exponential (14).

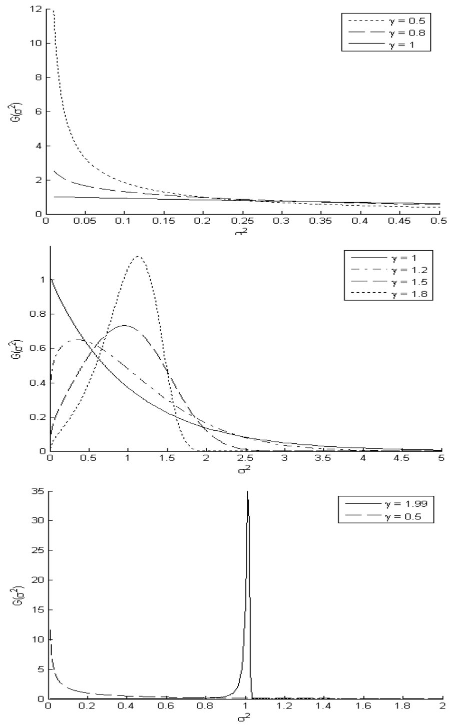

Figure 6. Gaussian Transforms of GGD

Unfortunately, the real part of the complex probability function diverges for periodical values of γ, which impedes on the computation of the transform through this method for $\gamma>2$. However, real data is mostly confined to $0<\gamma<2$ [6].

Generalized Cauchy Distribution (GCD)

The GCD probability density function is given by

$$p_x^C\left(x\mid\nu,b\right)=\frac{\Gamma\left(\nu+1/2\right)}{\sqrt{\pi}\Gamma(\nu)}\frac{b^{2\nu}}{\left(b^2+x^2\right)^{\nu+0.5}},$$

and it particularizes to the Cauchy distribution for $\nu=0.5$. Its Gaussian Transform can be computed through:

$$G_{X|\nu,b}\left(\sigma^2\right)=\frac{b^{2\nu}\Gamma(\nu+0.5)}{\sigma^2\sqrt{2\sigma^2}\Gamma(\nu)}\left(\mathcal{F}^{-1}\left(\left(b^2+i\omega\right)^{-\nu-0.5}\right)\right)_{t=\frac{1}{2\sigma^2}}.$$

Corresponding plots are given in Fig.7.

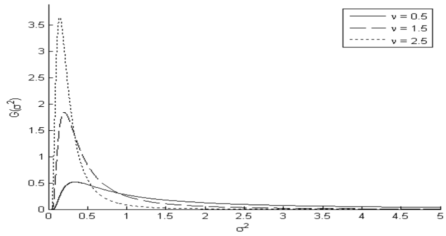

Figure 7. Gaussian Transforms of GCD

As a remark, the variance of the Cauchy distribution being infinite, its Gaussian Transform has infinite mean value (7). The GCD family possesses finite variance only for $\nu>1$.

4. CONCLUSION

We introduced in this paper the Gaussian Transform, which allows for the representation of symmetric distributions as an infinite Gaussian mixture. The scope of applicability of the Gaussian Transform is potentially very broad, from denoising and regularization to filtering, coding ,compression, watermarking etc. However, an extension of the concept to non-symmetric distribution would be required for some specific applications. Further investigation of the existence conditions, especially non-negativity, is also necessary. Finally, one would need adapted numerical packages (most likely based on existing Laplace and Fourier transform computational packages) for the computation of Gaussian Transforms of both analytically and numerically defined distributions.

ACKNOWLEDGMENT

This work is supported in part by the Swiss NCCR (National Centre of Research) IM2 (Interactive Multimodal Information Management, http://www.im2.ch) and the EU Network of Excellence Similar (http://www.similar.cc).
The authors thank Dr. Oleksiy Koval for careful reading and fruitful discussions.

REFERENCES

1. A. D. Poularikas, S. Seely, “Laplace Transforms. The Transforms and Applications Handbook: Second Edition”. Ed. Alexander D. Poularikas Boca Raton: CRC Press LLC, 2000
2. G. Arfken, Mathematical Methods for Physicists, 3rd ed. Orlando, FL: Academic Press, pp. 481-485, 1985.
3. G. Arfken, "Inverse Laplace Transformation." 15.12 in Mathematical Methods for Physicists, 3rd ed. Orlando, FL: Academic Press, pp. 853-861, 1985.
4. R. Wilson, “MGMM: Multiresolution Gaussian Mixture Models for Computer Vision”, in IEEE International Conference on Pattern Recognition, Barcelona, Spain, September 2000
5. A. Hjørungnes, J. M. Lervik, and T. A. Ramstad , “Entropy Coding of Composite Sources Modeled by Infinite Mixture Gaussian Distributions”, in Proc. 1996 IEEE Digital Signal Processing Workshop , Loen, Norway, September 1996
6. S.G. Mallat, “A Theory for Multiresolution Signal Decomposition: The Wavelet Representation”, in IEEE Transactions on Pattern Analysis and Machine Intelligence Volume 11 , Issue 7 (July 1989) p. 674-693.