ESANN 2012 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 25-27 April 2012, i6doc.com publ., ISBN 978-2-87419-049-0.
Available from http://www.i6doc.com/en/livre/?GCOI=28001100967420.

# Combined scattering for rotation invariant texture analysis

Laurent Sifre[1] and Stéphane Mallat[1,2] *

1- Ecole Polytechnique - Centre de Mathématiques Appliquées
Palaiseau - France

2- Institut des Hautes Etudes Scientifiques
Bures-sur-Yvette - France

**Abstract**. This paper introduces a combined scattering representation for texture classification, which is invariant to rotations and stable to deformations. A combined scattering is computed with two nested cascades of wavelet transforms and complex modulus, along spatial and rotation variables. Results are compared with state-of-the-art algorithms, with a nearest neighbor classifier.

## 1 Introduction

Texture classification has many applications from satellite to medical imagery. In these contexts, textures are typically *rotated* because of variations of the observer orientation, and their projection in the image plane undergoes *small deformations* due to 3D perspective effects.

Many methods [1] start by computing informative statistics with a *first computational layer* and then use a *second layer* to build rotation invariance. Second layers are designed with different strategies. *Registration or normalization* of rotation (or affine [2]) parameters keeps most of the information but typically suffers from instabilities, particularly when the texture is deformed. *Averaging* statistics (LBP[3], RI-LPQ[4]) along rotation is stable but loses all relative angular distribution of texture components. This angular distribution may be captured by computing a Fourier transform along the rotation parameter, whose modulus is rotation invariant, as in LBP-HF[5]. However, high frequency Fourier coefficients are known to be unstable in presence of deformations.

A texture $X(x)$ is modeled as a realization of a stationary process. A *representation* of the texture $X$ is a deterministic quantity that does not depend on the realization but only on the law of $X$ (e.g. it's autocorrelation matrix). Because of perspective effects, this texture may be deformed by $L_\tau X(x) = X(x - \tau(x))$ where $\tau(x)$ is modeled as an unknown stationary random process independant of $X(x)$, and it may be rotated by $r$ into $rX(x) = X(rx)$. A representation is stable to deformation if a small deformation (i.e. when $\sup_x \|\nabla\tau(x)\|$ is small) induces small changes in the representation. It is invariant to rotation if the representation does not change when $X$ is rotated.

This paper introduces a *combined expected scattering* representation and a *combined windowed scattering* estimator that keep most of the process information while being invariant to rotations and stable to deformations. A first

*This work has been founded by the french ANR 0126 01

ESANN 2012 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 25-27 April 2012, i6doc.com publ., ISBN 978-2-87419-049-0.
Available from http://www.i6doc.com/en/livre/?GCOI=28001100967420.

layer computes statistics of stationary textures that are informative, stable to deformations and non-rotation invariant. These coefficients are retransformed through a second layer to achieve rotation invariance while maintaining stability to deformations and most of the information.

The first layer is obtained by the scattering transform introduced in [6]. Scattering transform computes recursive co-occurence coefficients through a cascade of wavelet-modulus operators, along a convolutional network [7]. It has been applied to audio [8] and image [9] classification. The second layer is designed with a similar algorithm with convolutions computed *along the rotation parameter*. Angular information is scattered into different paths before being averaged along the rotation parameter. The resulting decomposition has the stability of averaging algorithms and the near completness properties of Fourier spectral approaches.

Expected scattering and windowed scattering are defined in Section 2. The combined scattering algorithm is presented in Section 3 and summarized in Figure (1). Section 4 shows the resulting improvements obtained for rotation invariant texture classification on the OUTEX 10 database [10], in comparison with state-of-the-art algorithms [3, 4, 5]. Softwares are available at `www.cmap.polytechnique.fr/scattering`.

## 2 Texture and spatial scattering

### 2.1 Modulus of complex wavelet

The frequency displacement induced by a deformation is proportionnal to the frequency, making Fourier analysis unstable to deformations at high frequencies [6]. A wavelet $\psi$ is a complex band-pass filter which is spatially well localized and thus stable to deformations. Its rotation by $r$ and dilation is written

$$\psi_\lambda(x) = 2^{2j}\psi(2^j r^{-1}x) \quad \text{with} \quad \lambda = 2^j r \ .$$

The wavelet-modulus coefficients of $X(x)$ are $U[\lambda]X(x) = |X * \psi_\lambda(x)|$. The mean and variance of such coefficients have been used by many algorithms [11, 12] to discriminate textures.

### 2.2 Spatial scattering

The full spatial variability of $U[\lambda]X(x)$ is not fully captured by its mean and variance. In [6, 9] this information is recovered by iteratively applying $U[\lambda]$ before evaluating the mean. The texture variability information is *scattered* into different *paths* $p = (\lambda_1, ..., \lambda_m)$ with the *scattering propagator* $U[p]$ defined by:

$$U[p]X(x) = U[\lambda_m]\dots U[\lambda_1]X(x) = \big|\big|\dots||X * \psi_{\lambda_1}| * \psi_{\lambda_2}|\dots\big| * \psi_{\lambda_m}(x)\big| \ .$$

The *expected scattering* is defined by

$$\bar{S}[p]X = \mathbb{E}\big(U[p]X(x)\big) = \mathbb{E}\Big(\big|\big|\dots||X * \psi_{\lambda_1}| * \psi_{\lambda_2}|\dots\big| * \psi_{\lambda_m}(x)\big|\Big) \ .$$

ESANN 2012 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 25-27 April 2012, i6doc.com publ., ISBN 978-2-87419-049-0.
Available from http://www.i6doc.com/en/livre/?GCOI=28001100967420.

It is proved in [6] that an expected scattering $\bar{S}$ is stable to elastic deformations. $\bar{S}$ is estimated from a single realization of $X$ with a *windowed scattering* $S_J$, which performs a spatial averaging on a domain whose width is proportional to $2^J$:

$$S_J[p]X(x) = U[p]X * \phi_J(x) = \big|\big| \dots ||X * \psi_{\lambda_1}| * \psi_{\lambda_2}| \dots \big| * \psi_{\lambda_m}\big| * \phi_J(x),$$

where $\phi_J = 2^{-2J}\phi(2^{-J}x)$ is an averaging kernel dilated by $2^J$. It is proved in [6] that if the autocovariance of $X$ is integrable, then $S_J$ is a consistent estimator of $\bar{S}$, and that, for suitable wavelets, $S_J$ preserves the mean square energy of $X$.

## 3 Combined scattering for rotation invariance

### 3.1 Scattering covariance with rotation

Rotation acts on the spatial variable $x$ but also on the wavelet parameter $\lambda$:

$$U[\lambda](rX)(x) = U[r\lambda]X(rx).$$

Taking an expected value proves that

$$\bar{S}[p](rX) = \bar{S}[rp]X,$$

where $rp = (r\lambda_1, r\lambda_2, ..., r\lambda_m)$ is called the *orbit* of $p$ along the rotation group when $r$ varies. To obtain a rotation invariant representation which is stable to deformations, we apply the same strategy as in Section 2 but along the rotation parameter.

### 3.2 Combined scattering

Wavelets $\widetilde{\psi}_{\widetilde{\lambda}}(r)$ along the rotation $r$ are defined as $2\pi$ periodic wavelets [6]. For any function $h(p)$ of a path variable, the $2\pi$ periodic convolution with $\widetilde{\psi}_{\widetilde{\lambda}}(r)$ is defined by:

$$h \circledast \widetilde{\psi}_{\widetilde{\lambda}}(p) = \int_r h(r^{-1}p)\, \widetilde{\psi}_{\widetilde{\lambda}}(r) dr.$$

We use $\circledast$ instead of $*$ to indicate that convolutions are done along rotations. In practice $\circledast$ is a periodic convolution defined on a small finite set. Cascading wavelet-modulus operators $\widetilde{U}[\widetilde{\lambda}]f = |f \circledast \widetilde{\psi}_{\widetilde{\lambda}}|$ along rotation parameter scatters the angle distribution information of the expected scattering along several combined paths $(\widetilde{p}, p) = (\widetilde{\lambda}_1, \dots, \widetilde{\lambda}_{\widetilde{m}}, \lambda_1, ..., \lambda_m)$ ,

$$\widetilde{U}[\widetilde{p}, p]X = \widetilde{U}[\widetilde{\lambda}_{\widetilde{m}}] \dots \widetilde{U}[\widetilde{\lambda}_1]\bar{S}[p]X = \big|\big| \dots ||\bar{S}X \circledast \widetilde{\psi}_{\widetilde{\lambda}_1}| \circledast \widetilde{\psi}_{\widetilde{\lambda}_2}| \dots \big| \circledast \widetilde{\psi}_{\widetilde{\lambda}_{\widetilde{m}}}[p]\big|$$

and rotation invariance is obtained by a final average along the rotation parameter at scale $2^L$.

$$\widetilde{\bar{S}}_L[\widetilde{p}, p]X = (\widetilde{U}[\widetilde{p}, p]X) \circledast \widetilde{\phi}_L[p] = \big|\big| \dots ||\bar{S}X \circledast \widetilde{\psi}_{\widetilde{\lambda}_1}| \circledast \widetilde{\psi}_{\widetilde{\lambda}_2}| \dots \big| \circledast \widetilde{\psi}_{\widetilde{\lambda}_{\widetilde{m}}}\big| \circledast \widetilde{\phi}_L[p],$$

ESANN 2012 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 25-27 April 2012, i6doc.com publ., ISBN 978-2-87419-049-0.
Available from http://www.i6doc.com/en/livre/?GCOI=28001100967420.

where $2^L$ is the rotation invariance scale. If $2^L = 1$ then $\widetilde{\phi}_L$ is constant and the representation is fully rotation invariant. Combined expected scattering $\widetilde{\bar{S}}_L$ is defined with an expected value operator $\mathbb{E}$ resulting from $\bar{S}$. It is estimated from a single realization $X$, by replacing $\bar{S}$ by $S_J$. This yields the combined windowed scattering:

$$\widetilde{S}_{L,J}[\widetilde{p}, p]X(x) = \big|\big| \ldots |S_J X \circledast \widetilde{\psi}_{\widetilde{\lambda}_1}| \ldots \big| \circledast \widetilde{\psi}_{\widetilde{\lambda}_{\widetilde{m}}}\big| \circledast \widetilde{\phi}_L[p](x) \ .$$

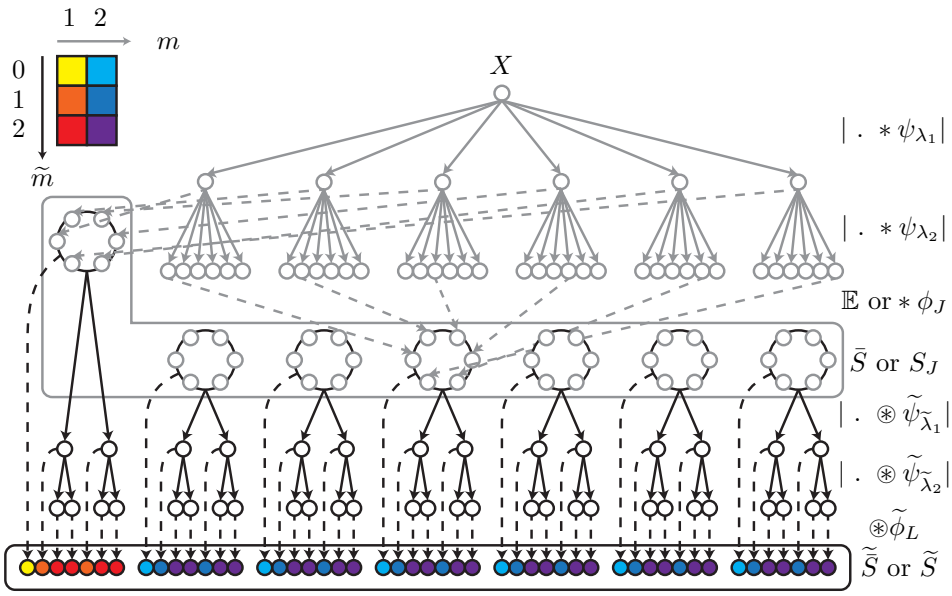



Fig. 1: Combined scattering architecture. First layer in grey, second layer in black. Spatial wavelet-modulus operators (grey arrows) are averaged (doted grey arrows), as in [9]. Outputs of the first layer are reorganized in different orbits (large black circles) of the action of the rotation on the representation. A second cascade of wavelet-modulus operators along the orbits (black arrows) split the angular information in several combined paths that are averaged (doted black arrows) along the rotation to achieve rotation invariance. Output nodes are colored with respect to the order of their corresponding paths.

The combined scattering algorithm is summarized in Figure 1. For an image of size $N$ the total computational complexity is $O(N \log N)$ and the resulting representation is much less than $N$ when $N$ is large. Let $m_{\max}$ and $\widetilde{m}_{\max}$ be the maximum length of $p$ and $\widetilde{p}$. In application, we choose $m_{\max} \leq 2$, $\widetilde{m}_{\max} \leq 2$ and we only compute paths $p = (2^{j_k} r_k)_k$ and $\tilde{p} = (2^{\tilde{j}_k})_k$ for decreasing frequencies $2^{j_{k+1}} \leq 2^{j_k}$ and $2^{\tilde{j}_{k+1}} \leq 2^{\tilde{j}_k}$, because they carry most of the scattering energy. Let us consider a textured image of $N = 2^{2J} = 2^{10} = 1024$ pixels ($J = 5$). For a total of $T = 8$ rotations, the spatial scattering represents this image with $TJ + T^2 J(J-1)/2 = 680$ coefficients. It corresponds to 85 orbits of 8

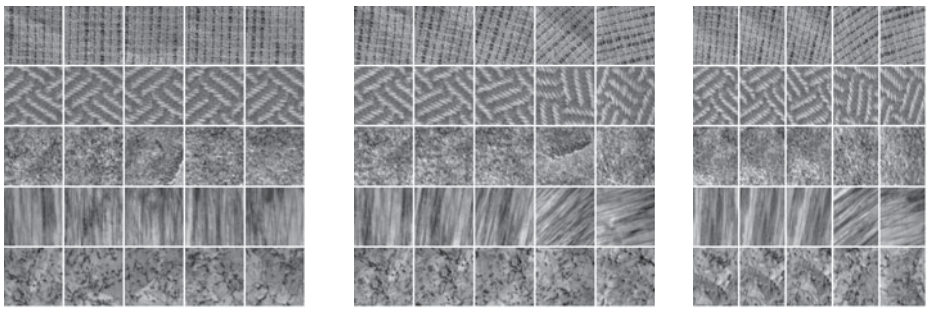
ESANN 2012 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 25-27 April 2012, i6doc.com publ., ISBN 978-2-87419-049-0.
Available from http://www.i6doc.com/en/livre/?GCOI=28001100967420.

coefficients each. The combined scattering keeps almost constant the number of coefficients representing each orbit. When $N = 2^{2J}$ is larger then the scattering representation $TJ + T^2J(J-1)/2$ becomes much smaller than $N$.

## 4 Texture Classification

Texture classification experiments are performed on the OUTEX10 database [10] (*rot* experiment). It contains 24 different texture classes. Each class has 20 training samples with a single orientation which is normalized to 0°. There are $24 \times 20 \times 8$ testing samples corresponding to 20 samples in each class that are rotated by $10°, 20°, \ldots, 90°$.

A second experiment (*rot-tilt*) simulates a perspective effect called tilt [13]. It is implemented with a gaussian blur with $\sigma = \sqrt{1.3^2 - 1}$ and a subsampling at intervals 1.3 *in the horizontal direction only*, for all images in the testing set. Images in the training set are cropped to keep the same image size in both sets. Figure 2 shows some training and testing samples from both experiments.



Training set Testing for *rot* Testing for *rot-tilt*

Fig. 2: A few samples of the databases used for experiments *rot* and *rot-tilt*

A nearest neigbor classifier is applied to the combined scattering $\widetilde{S}_{0,J}$ representation with several choices of maximum path length $m_{\max}, \tilde{m}_{\max} \leq 2$, and to other state-of-the-art descriptors for rotation invariant texture analysis. $L_2$ distance is used except for LBP-HF where the authors recommand [5] to use $L_1$. LBP[3] computes histograms of local binary patterns. Bins that correspond to rotated versions of the same pattern are merged, which leads to a loss of discriminability. LBP-HF[5] computes a Fourier transform modulus on the rotation parameter of LBP[3]. It thus maintains variability information along angles while achieving rotation invariance. RI-LPQ[4] computes windowed Fourier coefficients over a discrete set of frequencies distributed along circles. The phase is quantized to obtain a binary word on which a histogram is computed. As opposed to LBP and LBP-HF, RI-LPQ is robust to image blurring [4].

Results are presented in Table 1. The combined scattering achieves the best results with or without tilt distortions. The classification accuracy is improved

ESANN 2012 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 25-27 April 2012, i6doc.com publ., ISBN 978-2-87419-049-0.
Available from http://www.i6doc.com/en/livre/?GCOI=28001100967420.

| | *rot* | *rot-tilt* |
|---|---|---|
| $\text{LBP}^{riu2}/\text{VAR}_{(8,1)+(16,2)+(24,3)}$ $(r)$ | 97.7 | NC |
| $\text{LBP-HF}_{(8,1)+(16,2)+(24,3)}$ (c ) | 96.59 | 67.50 |
| RI-LPQ $(c)$ | 98.26 | 78.02 |
| $\widetilde{S}_{0,J}$, $m_{\max}, \widetilde{m}_{\max} = 1, 2$ | 96.72 | 81.61 |
| $\widetilde{S}_{0,J}$, $m_{\max}, \widetilde{m}_{\max} = 2, 0$ | 97.73 | 89.38 |
| $\widetilde{S}_{0,J}$, $m_{\max}, \widetilde{m}_{\max} = 2, 1$ | 98.62 | 92.89 |
| $\widetilde{S}_{0,J}$, $m_{\max}, \widetilde{m}_{\max} = 2, 2$ | **98.75** | **93.07** |

Table 1: Classification results $(r)$ is reported from papers, $(c)$ is obtained with authors' software [14].

when adding second order paths both in space and in rotation, with $m_{\max} = \tilde{m}_{\max} = 2$. For the *rot-tilt* experiment, the scattering brings an important improvement because it is stable to deformation, which is not the case for other rotation invariant representations [4, 5].

Future work will address more complex geometric invariance such as scaling, as well as supervised class-specific invariance.

## References

[1] J. Zhang. Brief review of invariant texture analysis methods. *Pattern Recognition*, 35(3):735–747, 2002.

[2] J. Zhang and T. Tan. Affine invariant classification and retrieval of texture images. *Pattern Recognition*, 36(3):657–664, 2003.

[3] T. Ojala, M. Pietikäinen and T. Mäenpää. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *PAMI*, 24(7):971 –987, 2002.

[4] V. Ojansivu, E. Rahtu, and J. Heikkila. Rotation invariant local phase quantization for blur insensitive texture analysis. *ICPR Proc.*, pages 1 –4, 2008.

[5] T. Ahonen, J Matas, C. He and M. Pietikäinen. Rotation invariant image description with local binary pattern histogram fourier features. *SCIA Proc.*, pages 61–70, 2009.

[6] S. Mallat. Group invariant scattering, http://arxiv.org/abs/1101.2286.

[7] Y. LeCun, K. Kavukvuoglu and C. Farabet. Convolutional networks and applications in vision. *ISCAS Proc.*, pages 253–256, 2010.

[8] J. Andén and S. Mallat. Multiscale scattering for audio classification. *ISMIR Proc.*, pages 657–662, 2011.

[9] J. Bruna and S. Mallat. Classification with scattering operators. In *CVPR*, pages 1561–1566, 2011.

[10] T. Ojala, T. Mäenpää, M. Pietikäinen, J. Viertola, J. Kyllönen and S. Huovinen. Outex - new framework for empirical evaluation of texture analysis algorithms. *ICPR Proc.*, pages 10701–, 2002.

[11] R. Porter and N. Canagarajah. Robust rotation-invariant texture classification: wavelet, gabor filter and gmrf based schemes. *VISP Proc.*, 144(3):180–188, 1997.

[12] M. Unser. Texture classification and segmentation using wavelet frames. *Image Processing, IEEE Trans.*, 4(11):1549–1560, 1995.

[13] G. Yu and J.M. Morel. A fully affine invariant image comparison method. *ICASSP Proc.*, pages 1597 –1600, 2009.

[14] http://www.cse.oulu.fi/mvg/downloads.