Combined scattering for rotation invariant texture analysis

Laurent Sifre and Stéphane Mallat

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 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 deformations.

At texture $X(x)$ is modeled as a fast stationary process. A representation of the texture is a deterministic quantity that does not depend on the realization but only on the law of $X(x)$ (e.g. its autocorrelation matrix).

Because of perspective effects, this texture may be deformed by $L_\tau X(x) = X(x - \tau(x))$ where $\tau(x)$ is a stationary random process independent 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 $\|\nabla \tau(x)\|$ is small) induces small changes in the representation. It is invariant to rotation if the representation does not change when $X(x)$ 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...
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]. Scat-
tering transform computes recursive co-occurrence 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 parame-

Angular information is scattered into different paths before being averaged
along the rotation parameter. A n 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 completeness 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 rota-
tion invariant texture classification on the OUTEX 10 database [10], in com-
parison 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

\[ \psi_\lambda(x) = 2^{2j} \psi(2^{j-1}) \] \[ \lambda = 2^j . \]

2.2 Spatial scattering

\[ U[ ]X( ) = U[\lambda_1]...U[\lambda_n]X( ) = [X * \psi_{\lambda_1}] * [X * \psi_{\lambda_2}] * ... * [X * \psi_{\lambda_n}] . \]

The expected scattering \( S[X] = \mathbb{E}(U[ ]X( )) = \mathbb{E}( [X * \psi_{\lambda_1}] * [X * \psi_{\lambda_2}] * ... * [X * \psi_{\lambda_n}] ) \).
where wavelet-modulus operators. In practice, a windowed scattering \( S_J \) serves as a localized counterpart of the singular value decomposition. This localized counterpart is essential for the treatment of data that is not uniformly distributed, such as signal analysis.

When \( r \) varies, to obtain a rotation-invariant representation which is stable to elastic deformations, we apply the same strategy as in Section 2 but along the rotation parameter. The windowed scattering \( S \) is estimated from a single realization of \( X \). It is proved in [6] that an expected scattering \( \bar{S} \) satisfies:

\[
\bar{S}_{\lambda} = \langle |X - \mu_Y| \psi_{\lambda} \rangle = \langle |X - \mu_Y| |X - \mu_Y| \psi_{\lambda} \psi_{\lambda} \rangle
\]

where \( \mu_Y \) is the mean of \( Y \) and \( \lambda \) a parameter at scale 2.

3 Combined scattering for rotation invariance

3.1 Scattering covariance with rotation

\( R \) a a e x a a \( \lambda \), \( \lambda_a \) a b \( \lambda \) a e d \( \lambda \) a \( \lambda \):

\[
U[\lambda](X)(\lambda) = U[\lambda]X(\lambda).
\]

\( T \) a a e d \( \lambda \) a \( \lambda \) a e d \( \lambda \) a \( \lambda \).

3.2 Combined scattering

\( W \) a a c e \( \psi_X^\lambda(\lambda) \) a a b \( \lambda \) a e d \( \lambda \) a b \( \lambda \):

\[
\bar{S}^\lambda = \langle |X - \mu_Y| \psi_{\lambda} \rangle = \langle |X - \mu_Y| |X - \mu_Y| \psi_{\lambda} \psi_{\lambda} \rangle
\]

We recall that for a \( \psi \), \( \psi_X^\lambda(\lambda) \) is a periodic convolution defined on a small finite set. Cascading \( \psi \) to indicate that convolutions are done along rotations.

\( X \) is a periodic convolution defined on a small finite set. Cascading \( \psi \) instead of \( \psi \) along the rotation parameter scatters the angle distribution information of the expected scattering along several combined paths.

S. o. m. b. \( p, p \) e \( \lambda \) \( \bar{S} \), a d \( \lambda \) a \( \lambda \) a e d \( \lambda \) a \( \lambda \).

S. o. m. b. \( p, p \) e \( \lambda \) \( \bar{S} \), a d \( \lambda \) a \( \lambda \) a e d \( \lambda \) a \( \lambda \).

\[
\bar{S}_{\lambda} = \langle |X - \mu_Y| \psi_{\lambda} \rangle = \langle |X - \mu_Y| |X - \mu_Y| \psi_{\lambda} \psi_{\lambda} \rangle
\]

where \( \mu_Y \) is the mean of \( Y \) and \( \lambda \) a parameter at scale 2.

3 Combined scattering for rotation invariance

3.1 Scattering covariance with rotation

\( R \) a a e x a a \( \lambda \), \( \lambda_a \) a b \( \lambda \) a e d \( \lambda \) a \( \lambda \):

\[
U[\lambda](X)(\lambda) = U[\lambda]X(\lambda).
\]

\( T \) a a e d \( \lambda \) a \( \lambda \) a e d \( \lambda \) a \( \lambda \).

3.2 Combined scattering

\( W \) a a c e \( \psi_X^\lambda(\lambda) \) a a b \( \lambda \) a e d \( \lambda \) a b \( \lambda \):

\[
\bar{S}^\lambda = \langle |X - \mu_Y| \psi_{\lambda} \rangle = \langle |X - \mu_Y| |X - \mu_Y| \psi_{\lambda} \psi_{\lambda} \rangle
\]

We recall that for a \( \psi \), \( \psi_X^\lambda(\lambda) \) is a periodic convolution defined on a small finite set. Cascading \( \psi \) to indicate that convolutions are done along rotations.

\( X \) is a periodic convolution defined on a small finite set. Cascading \( \psi \) instead of \( \psi \) along the rotation parameter scatters the angle distribution information of the expected scattering along several combined paths.
where $L$ is the rotation invariance scale. If $L = 1$, then $\tilde{\phi}_L$ is constant and the representation is fully rotation invariant. Combined expected scattering $\bar{S}_L$ is defined with an expected value operator $E$ resulting from $\bar{S}_L$. It is estimated from a single realization $X$, by replacing $\bar{S}_L$ by $S$. This yields the combined windowed scattering:

$$\tilde{S}_{L,J}[\tilde{\psi}_L] X = \cdots \ast \bar{S}_L \ast \tilde{\psi}_L \ast \tilde{\lambda}_1 \ast \cdots \ast \tilde{\phi}_L \ast \mathbf{X} \ast \tilde{\psi}_L \ast \tilde{\lambda}_2 \ast \cdots \ast \tilde{\phi}_L \ast \mathbf{X}.\) 

**Fig. 1:** Combined scattering architecture. First layer in grey, second layer in black. Spatial wavelet-modulus operators (grey arrows) are averaged (dotted 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. As second cascade of wavelet-modulus operators along the orbits (black arrows) split the angular information in several combined paths that are averaged (dotted 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 $m_{\text{max}}$ and $\tilde{m}_{\text{max}}$ be the maximum length of $p$ and $\tilde{p}$, respectively. In an application, we choose $m_{\text{max}} \leq 2$, $\tilde{m}_{\text{max}} \leq 2$, and we only compute paths $p = (2^j k)$ and $\tilde{p} = (2^j \tilde{k})$ because they carry most of the scattering energy.

Let us consider a textured image of $N = 2^J = 2^{10} = 1024$ pixels ($J = 5$). For a total of $T = 8$ rotations, the spatial scattering represents this image with $T J + T^2 (J - 1)/2 = 680$ coefficients. It corresponds to $85$ orbits of $8$.
coefficients each. The combined scattering keeps almost constant the number of coefficients representing each orbit. When \( N = 2 \) the scattering representation becomes smaller than \( N \).

4 Texture Classification

Texture classification experiments are performed on the OUTEX10 database \([10]\) (rot and rot-tilt). 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 sample sine each class that are rotated by 10°, 20°, ..., 90°.

As a second experiment (rot-tilt) simulates a perspective effect \([13]\). 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. rotate a image at each angle.

Figure 2 shows some training and testing samples from both experiments.

![Figure 2: A few samples of the databases used for experiments](image)

An areas within a class is applied to the combined scattering \( \tilde{S}_0, J \) representation with several choices of maximum path length \( m_{\text{max}} \), \( \tilde{m}_{\text{max}} \leq 2 \), 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 recommend \( L_1 \). LBP \([3]\) computes histograms of local binary patterns. Bin 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.
Table 1: Classification results ($r_i$) reported from papers, ($c_i$) obtained with authors' software [14].

<table>
<thead>
<tr>
<th>Method</th>
<th>rot</th>
<th>rot-tilt</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBP$^{t_{rot}}$/VAR$^{(8,1)+(16,2)+(24,3)} (c)</td>
<td>97.7</td>
<td>NC</td>
</tr>
<tr>
<td>LBP-HF$^{(8,1)+(16,2)+(24,3)} (c)$</td>
<td>96.59</td>
<td>67.50</td>
</tr>
<tr>
<td>RI-LPQ (c)</td>
<td>98.26</td>
<td>78.02</td>
</tr>
<tr>
<td>$\tilde{S}_{0,J_i}$ max. $\sim$ max = 1, 2</td>
<td>96.72</td>
<td>81.61</td>
</tr>
<tr>
<td>$\tilde{S}_{0,J_i}$ max. $\sim$ max = 2, 0</td>
<td>97.73</td>
<td>89.38</td>
</tr>
<tr>
<td>$\tilde{S}_{0,J_i}$ max. $\sim$ max = 2, 1</td>
<td>98.62</td>
<td>92.89</td>
</tr>
<tr>
<td>$\tilde{S}_{0,J_i}$ max. $\sim$ max = 2, 2</td>
<td>98.75</td>
<td>93.07</td>
</tr>
</tbody>
</table>

References