Improved Cross-correlation for Template Matching on the Laplacian Pyramid

Bonmassar, Giorgio

Boston University Center for Adaptive Systems and Department of Cognitive and Neural Systems

http://hdl.handle.net/2144/2343

Boston University
Improved Cross-correlation for Template Matching on the Laplacian Pyramid

Giorgio Bonmassar and Eric L. Schwartz

February 1998

Technical Report CAS/CNS-98-009

Permission to copy without fee all or part of this material is granted provided that: 1. The copies are not made or distributed for direct commercial advantage; 2. the report title, author, document number, and release date appear, and notice is given that copying is by permission of the BOSTON UNIVERSITY CENTER FOR ADAPTIVE SYSTEMS AND DEPARTMENT OF COGNITIVE AND NEURAL SYSTEMS. To copy otherwise, or to republish, requires a fee and / or special permission.

Copyright © 1998

Boston University Center for Adaptive Systems and
department of Cognitive and Neural Systems
677 Beacon Street
Boston, MA 02215
Improved Cross-correlation for Template Matching on the Laplacian Pyramid

Giorgio Bonmassar

Department of Biomedical Engineering, Boston University, 44 Commonwealth Street, Boston, MA 02215, USA. Email: giorgio@bu.edu

Eric L. Schwartz

Department of Cognitive and Neural Systems, Boston University, 677 Beacon Street, Boston, MA 02215, USA.

Abstract

Template matching via cross-correlation on Laplacian pyramid image architectures has been traditionally performed in a "coarse" to "fine" fashion. In the present paper, we show that by computing cross-correlation within each level of the pyramid independently, and considering the sum across (expanded) levels, a significant improvement in Peak to Correlation Energy (PCE) [9] is obtained. This result is illustrated with a number of numerical examples.

Key words: Correlation, Laplacian Pyramid, Template Matching, Pyramid Architectures.

The image-pyramid format [3] has been used in the past decade as a multi-resolution image format for a wide variety of applications (see for review [7,4]) which include image enhancement, pattern recognition, texture and motion analysis. Dyer [6] describes a set of possible strategies for performing matching directly on the multi-scale pyramidal representation. Dyer’s approach is top-down: peaks resulting from matches in a coarse representation are considered first, followed by finer scale windows which are used to verify the presence of the object. Many different approaches have been introduced so far to take advantage of the pyramidal architecture [1,11,12,10], but none specifically address the issue of how to improve the peak to correlation Energy (PCE) [9].

1 Following [5], we use template matching to describe a technique used in order to decide if a previously specified template is present in an image. Cross-correlation is adopted here as a measure of similarity between template and a region of the image edge maps.
of the cross-correlation. In the present paper we introduce a modified version of the classical cross-correlation algorithm especially designed for use on the Laplacian pyramid, and which, in principle, can be applied to other pyramid architectures, as well.

1 The pyramidal-correlation.

We propose the following new pyramidal-correlation operator ($o_{\text{pyr}}$):

$$i(x, y) \circ_{\text{pyr}} t(x, y) = \sum_{l=0}^{n_l-1} i_l(x, y) \circ t_l(x, y)$$  \hspace{1cm} (1)

where $n_l$ is the number of levels used in the pyramidal-correlation (see Fig. 1).

The basic idea is to sum correlations ($\circ$) computed between homologous levels of the pyramid. The single correlations can be efficiently performed using the FFT directly on the representation of the reduced levels\(^2\). Our motivation for eq. (1) is it eliminates cross-band correlation noise. Therefore we expect an improvement in the peak to correlation energy (PCE) [9]. This is shown, to be correct, both experimentally and analytically in the next section.

We examine the correlation obtained by applying a matched filter to an input image $i(x, y)$ which contains a template $t(x, y)$ in presence of additive uncorrelated white noise $n(x, y)$: $I(k, h) = T(k, h) + N(k, h)$ (we use capital letters for frequency). A convenient definition of PCE, which measures the sharpness of the peak (supposed here to be in the center) in the cross-correlation plane, is the following [9]:

$$PCE_{\text{corr}} = \frac{|c(0, 0)|^2}{E_c} = \frac{\beta \left| \int_{-\infty}^{+\infty} T^*(k, h) I(k, h) \, dk \, dh \right|^2}{\int_{-\infty}^{+\infty} \|T(k, h)\|^2 \|I(k, h)\|^2 \, dk \, dh}$$  \hspace{1cm} (2)

where $c(x, y)$ is the cross-correlation function, and $\beta$ is a constant\(^3\). $E_c$ is the cross-correlation energy: $E_c = \int_{-\infty}^{+\infty} |c(x, y)|^2 \, dx \, dy$. In the case of the pyramidal-

\(^2\) Zero-padding must be introduced if sizes of image and template differ (most common case), or if sizes are not a power of two.

\(^3\) The value of the constant $\beta$ depends on the normalization term used in the Fourier Transform, see [2].
correlation operator, defined in eq.(1), the PCE is:

\[
PCE_{pyr} = \frac{|c_{pyr}(0,0)|^2}{E_{pyr}} = \beta \frac{\sum_{i=0}^{n-1} \sum_{h=0}^{\infty} \|I^i(k,h)\| \|I^h(k,h)\| \, dk \, dh}{\sum_{i=0}^{n-1} \sum_{h=0}^{\infty} \|I^i(k,h)\|^2 \|I^h(k,h)\|^2 \, dk \, dh}
\]  

(3)

We now establish conditions for a better PCE in the pyramidal-correlation with respect to the traditional cross-correlation.

**Theorem 1** Define the following positive definite quantities \(\delta(i,j)\) by:

\[
|c_{pyr}(i,j)|^2 = \delta(i,j) |c(i,j)|^2
\]  

(4)

where \(i = 0, \ldots, N-1\), \(j = 0, \ldots, M-1\), and both \(c(i,j)\) and \(c_{pyr}(i,j)\) are \(N \times M\) matrices. Then,

\[
PCE_{pyr} \geq PCE_{xcorr}
\]  

(5)

if the following inequality holds:\(^4\)

\[
\delta(0,0) \geq \delta(i,j)
\]  

(7)

**Proof:**

First, the pyramidal-correlation lacks the following cross-band frequency terms:

\[
C_c(k,h) = \sum_{m=0}^{n-1} \sum_{n=0}^{m-1} T^{ma}(k,h) I^n(k,h)
\]  

(8)

We can rewrite eq.(3), introducing \(\delta(i,j)\) according to eq.(4), and dividing both numerator and denominator by \(\delta(0,0)\):

\[
PCE_{pyr} = \frac{|c(0,0)|^2}{|c(0,0)|^2 + \frac{\delta(1,0)}{\delta(0,0)} |c(1,0)|^2 + \frac{\delta(2,0)}{\delta(0,0)} |c(2,0)|^2 + \ldots}
\]  

(9)

\(^4\) The inequality is true because the pyramidal correlation does not contain the cross-terms expressed in eq.(8), which are in general noise terms. Furthermore, in the traditional cross-correlation, as the template shifts, the coefficient:

\[
|c(x,y)|^2 = \left| \int \int t(x'+x, y'+y) \left( t(x',y') + n(x',y') \right)^* \, dx' \, dy' \right|^2
\]  

(6)

will decrease, given the presence of uncorrelated white noise. This decrease is even greater for \(|c_{pyr}(x,y)|^2\) since the cross-terms increase as the image tends to become pure white noise.
Due to the missing terms in eq.(8), \( E_c \geq E_{pyr} \) or equivalently:

\[
\begin{align*}
|c(0,0)|^2 + |c(1,0)|^2 + |c(2,0)|^2 + \ldots & \geq \delta(0,0) |c(0,0)|^2 + \\
& + \delta(1,0) |c(1,0)|^2 + \delta(2,0) |c(2,0)|^2 + \ldots \quad (10)
\end{align*}
\]

Following from eq.(7) (i.e., \( \frac{\delta(i,j)}{\delta(0,0)} \leq 1, \forall i, j \)):

\[
\begin{align*}
|c(0,0)|^2 + |c(1,0)|^2 + |c(2,0)|^2 + \ldots & \geq |c(0,0)|^2 + \\
& + \frac{\delta(1,0)}{\delta(0,0)} |c(1,0)|^2 + \\
& + \frac{\delta(2,0)}{\delta(0,0)} |c(2,0)|^2 + \ldots \quad (11)
\end{align*}
\]

Combining the above inequality with eq.(9) and comparing it with eq.(2):

\[ PCE_{pyr} \geq PCE_{xcorr} \square. \]

In practice the above inequality is a strong one, leading to a considerable improvement: nearly an order of magnitude in PCE (see section on numerical examples).

2 Numerical examples.

The condition of Theorem 1 expressed in eq.(7) is supported by all our numerical simulations taken on a large database of images (see Table 1 for some examples, where the images are taken from MATLAB and from the Standard Images database of the University of East Anglia (UK)\(^5\)).

<table>
<thead>
<tr>
<th></th>
<th>Boats</th>
<th>Photographer</th>
<th>Mandrill</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highest ( \delta )</td>
<td>( \delta(0,0) = 1.9 \cdot 10^{-4} )</td>
<td>( \delta(0,0) = 3.4 \cdot 10^{-2} )</td>
<td>( \delta(0,0) = 8.7 \cdot 10^{-4} )</td>
</tr>
<tr>
<td>Second highest ( \delta )</td>
<td>( \delta(255,0) = 6.7 \cdot 10^{-5} )</td>
<td>( \delta(255,0) = 1.2 \cdot 10^{-3} )</td>
<td>( \delta(0,255) = 1.9 \cdot 10^{-5} )</td>
</tr>
<tr>
<td>( PCE_{xcorr} )</td>
<td>( 2.1 \cdot 10^{-5} )</td>
<td>( 2.0 \cdot 10^{-5} )</td>
<td>( 1.7 \cdot 10^{-5} )</td>
</tr>
<tr>
<td>( PCE_{pyr} )</td>
<td>( 6.1 \cdot 10^{-2} )</td>
<td>( 2.9 \cdot 10^{-1} )</td>
<td>( 2.7 \cdot 10^{-1} )</td>
</tr>
</tbody>
</table>

Table 1: Value and position of the first two peaks of fractionals "\( \delta \)", and PCEs, for a set of noiseless \((256 \times 256)\) images (autocorrelation). Note that the highest peak is always \( \delta(0,0) \).

We now numerically demonstrate the advantages of the pyramidal-correlation over the traditional one through simulations using white noise added to image samples. Table 2 shows peak height (PH), PCE of traditional cross-correlation, pyramidal-correlation and cross-correlation energies ($E_c$). These values are obtained by averaging one hundred values of PH, PCE and $E_c$ of (the cross-correlation) superimposed Gaussian White Noise with $\sigma = 0.004$ of the DOG (Difference of Gaussians) image of Lenna (256 x 256).

<table>
<thead>
<tr>
<th></th>
<th>PH</th>
<th>PCE ($10^{-3}$)</th>
<th>$E_c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional</td>
<td>29.75</td>
<td>115.26</td>
<td>76775.83</td>
</tr>
<tr>
<td>Pyramidal</td>
<td>12.37</td>
<td>725.36</td>
<td>2108.53</td>
</tr>
</tbody>
</table>

Table 2: Performance measure of the pyramidal-correlation.

3 Conclusion.

The pyramidal-correlation has lower energy, as expected, due to the absence of cross terms at different bands (see eq.(8)). This is the reason for the smaller pyramidal-correlation peaks. The most significant result is the six-times higher PCE of the new correlation operator proposed in eq.(1). Pyramidal cross-correlation, as seen in Fig. 2, is qualitatively much sharper and less noisy than the traditional cross-correlation, and the PCE difference is 7 dB.

<< INSERT FIGURE 2 HERE. >>

Fig. 1, shows the PCE as function of the number of levels used, starting from levels zero and increasing to a cumulative level $nl - 1$ (the image was a 256 x 256 image of Lenna), and as a function of noise.

References


Fig. 1. Plots of PCE as a function of number of levels (nl) for the pyramidal-correlation, summing-up subsequent levels. From top left to bottom right, the Gaussian white noise parameter was: $\sigma_1 = 0$, $\sigma_1 = 0.05$, $\sigma_1 = 0.1$, $\sigma_1 = 0.5$. This figures shows the noise-robustness of the pyramidal-correlation as a function of nl. As a result these examples exhibit a good PCE just by computing the first three levels of the pyramidal-correlation.
Fig. 2. The result of applying the pyramidal-correlation to a (256 x 256) image of Lenna with a (64 x 64) detail-image of her hat. The top plot shows matching peak with the traditional cross-correlation algorithm and the bottom plot shows the pyramidal-correlation. Clearly visible is the sharper peak and lower background noise in the bottom plot compared to the top one. A Difference of Gaussian (DOG) filter was applied to image and template, with mask size of 7, and a 1.6 ratio between the standard deviations of the excitatory and inhibitory Gaussians.