M4.1

Proceedings of 1989 IEEE International Conference on Acoustics, Speech, and Signal Processing, Glasgow, Scotland, May 1989.

MORPHOLOGICAL CORRELATION AND MEAN ABSOLUTE ERROR CRITERIA

Petros Maragos

Division of Applied Sciences, Harvard University, Cambridge, MA 02138.

ABSTRACT. In this paper the mean absolute error criterion for signal matching/detection is linked with a morphological signal correlation (a sum of minima). Several properties of this nonlinear correlation are investigated, its performance for signal detection is compared to the classical (sum of products) linear correlation, and its statistical form is calculated for speckle patterns.

1 Signal Correlations and ℓ_p Norms

Consider two real-valued d-dimensional ($d = 1, 2, \ldots$) discrete signals represented by the sequences $f(n)$ and $g(n)$, $n \in \mathbf{Z}^d$. For simplicity, assume temporarily that g is a signal pattern to be found in f. To find which shifted version of g "best" matches f a standard approach has been to search for the shift lag k that minimizes the *mean squared error* (MSE)

$$E_2(k) = \sum_{n \in W} [f(n+k) - g(n)]^2$$

over some subset W of $\mathbf{Z}^d$. Since $(a-b)^2 = a^2 + b^2 - 2ab$ for any reals a, b, under certain assumptions this matching criterion is equivalent to maximizing the *linear cross-correlation* between f and g:

$$\gamma_{fg}(k) = \sum_{n \in W} f(n+k) g(n)$$

Such ideas have provided the foundations for many decades of research in matched filtering and signal detection. The popularity of the MSE criterion is mainly based on its mathematical tractability. From a statistical viewpoint one could also claim that this approach is optimal if one of the signals f or g is corrupted by additive Gaussian noise distributions. But the perfect Gaussian assumption for real data is a myth, and statisticians have found that even for slight deviations from the Gaussian assumption other error criteria are more robust. One such criterion is the *mean absolute error* (MAE)

$$E_1(k) = \sum_{n \in W} |f(n+k) - g(n)|.$$

This MAE matching criterion has been applied to speech signals for pitch period detection [1] and to images for template matching. MAE criteria have also been applied to solving optimization problems in rank order filtering [2].

What appears to have not been done is linking the MAE matching criterion with some signal correlation. This paper

This work was supported by the National Science Foundation under Grant MIPS-86-58150 with matching funds from Bellcore and Xerox, and in part by the ARO under Grant DAALO3-86-K-0171.

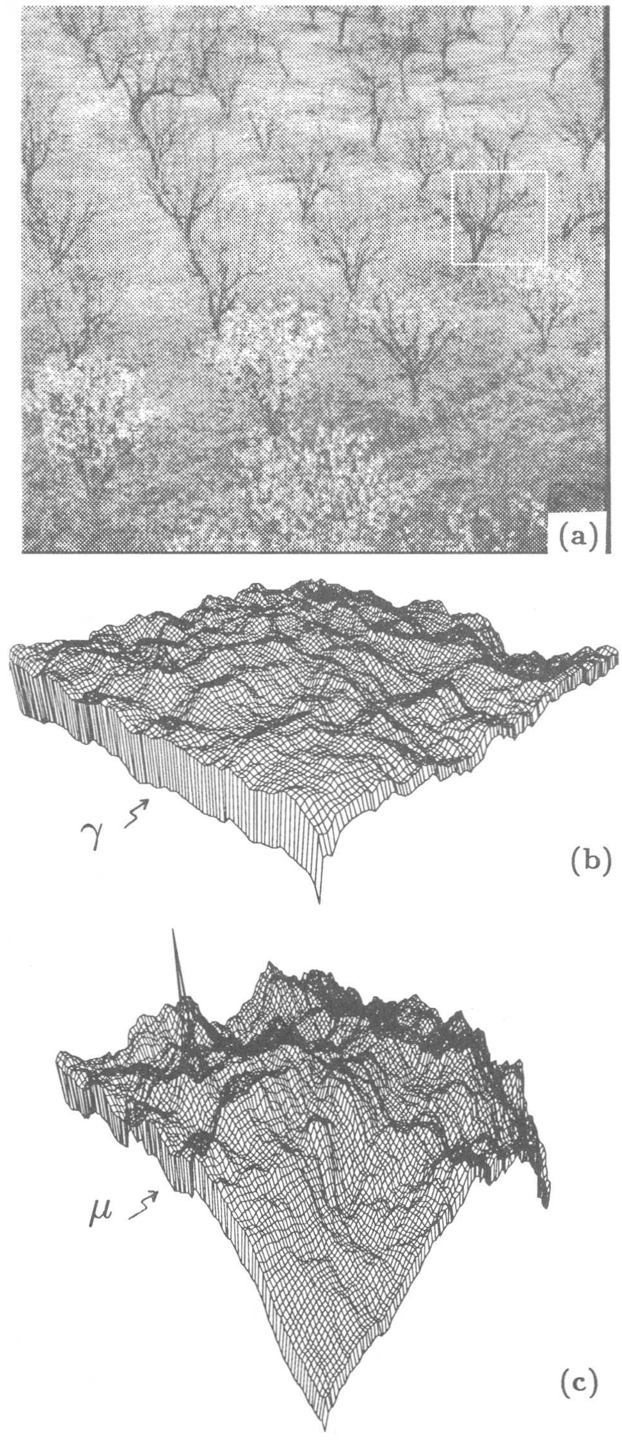


Figure 1. (a) An image f and a template g (inside the window). Normalized correlations: (b) Linear (γ). (c) Morphological (μ).

provides such a missing link. Namely, since $|a - b| = a + b - 2\min(a, b)$, under certain assumptions, minimizing E_1 is equivalent to maximizing the *nonlinear* cross-correlation

$$\mu_{fg}(k) = \sum_{n \in W} \min[f(n+k), g(n)].$$

If $W = \mathbf{Z}^d$, then the E_2 and E_1 error measures become respectively the ℓ_2 (squared) and ℓ_1 norms of the error signal $f - g$. In practice, W is often a *window*, i.e., a finite subset of $\mathbf{Z}^d$; then, in most cases, we can still view the errors E_2, E_1 as $\ell_{p=1,2}$ norms and the correlations γ, μ as defined for $-\infty < n, k < \infty$ if we first window f or g or both.

In [3] we showed that maximizing $\mu(k)$ is optimum for image matching or object detection under a variety of MAE matching criteria. In this paper we discuss some deterministic and statistical properties and applications of this nonlinear correlation, extending our work in [3]. To motivate our next discussion consider a 2-D signal f representing the image of Fig. 1a, and let g represent the windowed orchard (an image template). Fig. 1b shows $\gamma_{fg}(k)$ normalized by the product of the rms value of g and the local rms value of f. Fig. 1c shows $\mu_{fg}(k)$ normalized by the average of the area of g and the local area of f. As observed in Fig. 1, the detection of g in f is indicated by a much sharper peak in $\mu_{fg}(k)$ than in $\gamma_{fg}(k)$. In addition, the nonlinear correlation μ (a sum of minima) is faster than the linear (sum of products) correlation γ.

2 Morphological Correlation

We call $\mu_{fg}(k) = \sum_{n=-\infty}^{\infty} \min[f(n+k), g(n)]$ a *morphological* cross-correlation and $\mu_{ff}(k)$ a morphological autocorrelation of f, because it is directly related to a basic operation of mathematical morphology [4]. Specifically, $\mu_{ff}(k)$ is equal to the area under the signal obtained by *eroding* the function f by a 2-point structuring element $\{0, k\}$. (The relation between linear autocorrelation of binary images and erosion was used in [4] for structural image analysis via geometric probabilities.)

First note that

$$\mu_{fg}(k) \le \frac{A(f) + A(g)}{2} \quad \text{and} \quad \mu_{ff}(k) \le \mu_{ff}(0),$$

where $A(f) = \sum_n f(n)$ is the *area* under f.

Further, to relate γ and μ we assume that both f and g are *nonnegative* signals (an assumption easily met, e.g. by adding a dc-offset to the signals or when dealing with image signals), and we consider their *binary threshold signals*

$$f_a(n) = \begin{cases} 1, & f(n) \ge a \\ 0, & f(n) < a \end{cases}$$

where a spans *all* the continuous range of f. Then f can be represented exactly by its threshold signals since

$$f(n) = \int_0^\infty f_a(n)da.$$

Then we showed in [3] that the morphological correlation between f and g is the average over all amplitudes a of the binary correlations between f_a and g_a:

$$\mu_{fg}(k) = \int_0^\infty \gamma_{f_a g_a}(k)da = \int_0^\infty \mu_{f_a g_a}(k)da. \quad (1)$$

Note that for *binary* signals (but not otherwise) the morphological and linear correlation coincide. Result (1) has a counterpart result for the ℓ_p matching error norms $||f-g||_p$, $p = 1, 2$; i.e., as shown in [3],

$$||f-g||_1 = \int_0^\infty ||f_a - g_a||_1 da = \int_0^\infty (||f_a - g_a||_2)^2 da.$$

If $f(n) \overset{\mathcal{F}}{\longleftrightarrow} F(\omega) = \sum_n f(n)e^{-j\omega n}$ denotes a Fourier transform pair, it is well-known that the Fourier transform of $\gamma_{ff}(k)$ is the *energy spectral density* (or spectrum) $\Gamma(\omega) = |F(\omega)|^2$. Next we show something similar for the Fourier transform, $M(\omega)$, of the morphological correlation $\mu_{ff}(k)$. Namely, if $f_a(n) \overset{\mathcal{F}}{\longleftrightarrow} F_a(\omega)$ and $g_a(n) \overset{\mathcal{F}}{\longleftrightarrow} G_a(\omega)$, then $\mu_{fg}(k) \overset{\mathcal{F}}{\longleftrightarrow} \int_0^\infty F_a(\omega)G_a(-\omega)da$. This implies that

$$M(\omega) = \sum_{k=-\infty}^{\infty} \mu_{ff}(k)e^{-j\omega k} = \int_0^\infty |F_a(\omega)|^2 da. \quad (2)$$

Observe that $\frac{1}{2\pi}\int_{-\pi}^{\pi} M(\omega)d\omega = \mu_{ff}(0) = \sum_n f(n) = A(f)$. Hence, in a very small frequency zone $\Delta\omega$ where $M(\omega)$ remains approximately constant, we have that $M(\omega)\Delta\omega \approx \Delta A(f)$. This implies that $M(\omega)$ measures how the area of $f(n)$ distributes over frequency. Therefore we call $M(\omega)$ an *area spectral density* (or area spectrum) and, as (2) implies, it is the average (over all amplitudes a) of the energy spectra of the binary signals $f_a(n)$. Since $\mu_{ff}(k)$ is a real even signal, $M(\omega)$ is also real and even, but not necessarily nonnegative for signals f assuming both positive and negative values.

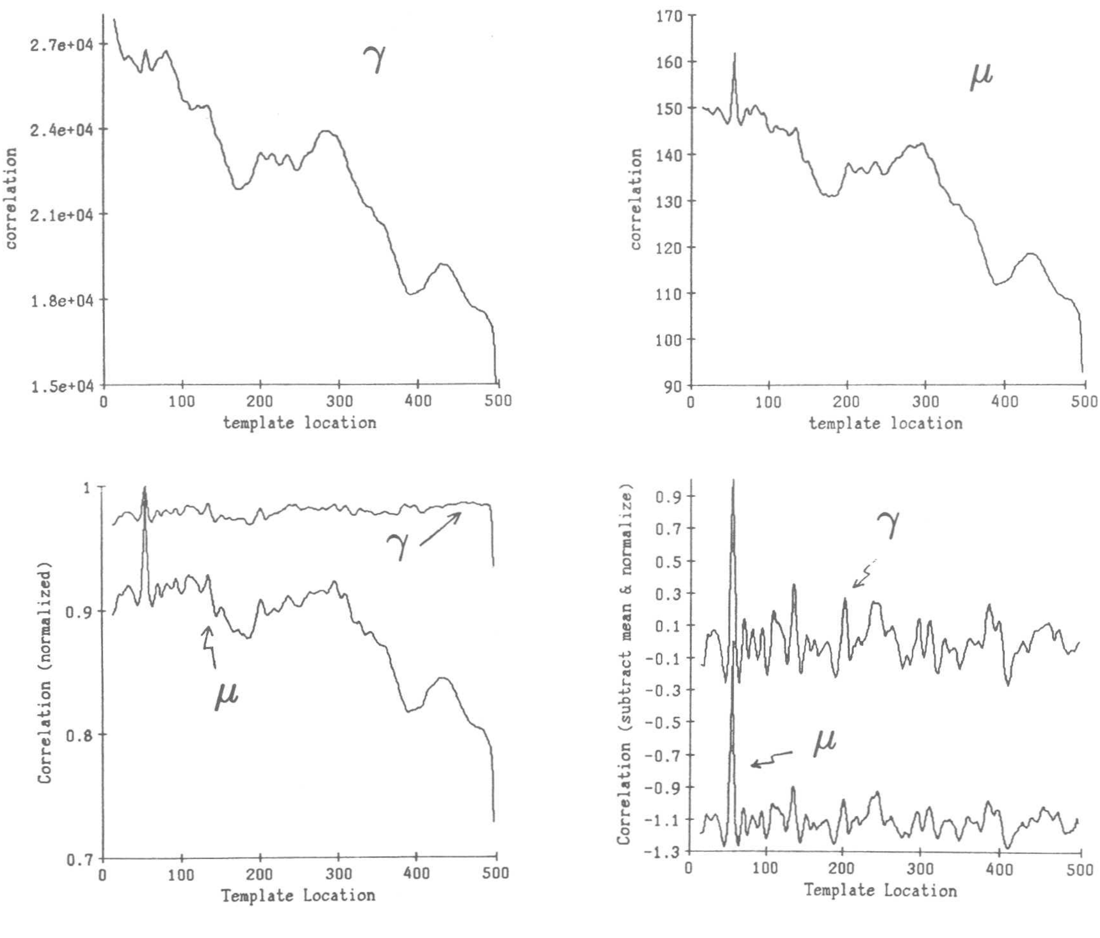
Experiments: The observation from Fig. 1 that morphological correlation yields *sharper peaks* than the linear correlation for signal matching/detection has been experimentally verified in many of our experiments both with *images* as well as with voiced *speech* signals. To see this clearly, Fig. 2 reports a series of 3 experiments for detecting in an image f a template g which is a part of f: A) The windowed (i.e., short-space) cross-correlations $\gamma_W = (1/|W|)\sum_{n \in W} f(n+k)g(n)$ and $\mu_W = (1/|W|)\sum_{n \in W} \min[f(n+k), g(n)]$ are shown in Figs. 2b,c, where W is the domain of g and $|W|$ is the number of pixels in g. B) The *normalized* correlations are shown in Fig. 2d. $\gamma_W(k)$ is normalized by dividing it with the product of the rms value of g and the local rms value of f, whereas $\mu_W(k)$ is normalized by dividing it with the average of the mean of g and the local mean of f. The range of both of these normalized correlations is $[0, 1]$. C) Fig. 2e shows the windowed morphological and linear cross-correlations among f and g after their means have been subtracted and they have been normalized by dividing them with their standard deviations. The correlations are further normalized so that both have a dynamic range of the same length, i.e, [-1,1] for the linear and [-2,0] for the morphological.

Which experiment (A, B, or C) is more meaningful for signal detection depends of course on the specific application. We generally observe, however, in the experiments A and B that the morphological correlation (either its direct or normalized version) has two practical advantages over the linear correlation: it yields sharper matching peaks and is faster since it is a sum of minima, as opposed to a sum of products for the linear. The sharper peaks of the morphological correlation were also observed when we repeated the same experiments for signals corrupted with impulse

Figure 2. (a) An image f (512 pixels in width) and a 30×30-pixel teplate g (inside the window). The horizontal line shows the direction along which cross-correlations of f and g were computed. (b) Linear correlation (γ). (c) Morphological correlation (μ). (d) Normalized correlations. (e) Correlations among f and g after their means were subtracted and their values were normalized by dividing them with the respective standard deviations.

(a)

(b) (c) (d) (e)



noise. In experiment C both correlations have similar computational complexity and performance in terms of signal detection. However, they are optimal under different error criteria, which are compatible with different ℓ_p norms or related noise distributions.

To explain the sharper peaks for the morphological correlation we provide next a theoretical analysis, which is similar to the approach followed in [1] to show that the MAE $E_1(k)$ gives sharper *valleys* for pitch detection in speech than the peaks of the linear correlation $\gamma_W(k)$. We assume that f and g are (locally) realizations of the same random sequence, which contains some quasi-periodic structure (the periodicity here refers to the repetitive occurrence of g in f, e.g. as in Fig. 2). Thus in the definitions of γ_W and μ_W we replace g with f. From Cauchy-Schwartz's inequality it can be shown that

$$\frac{1}{|W|}\sum_{n\in W}|f(n+k)-f(n)| \leq \sqrt{\frac{1}{|W|}\sum_{n\in W}|f(n+k)-f(n)|^2}$$

which implies that

$$\begin{array}{c}\frac{1}{|W|}\sum_{n\in W}[f(n+k)+f(n)]-2\mu_W(k)= \\ r(k)\sqrt{\frac{1}{|W|}\sum_{n\in W}[f^2(n+k)+f^2(n)]-2\gamma_W(k)}\end{array},$$

where $0\leq r(k)\leq 1$. Assuming also that f is (locally) *stationary* yields that $\sum_{n\in W}f(n+k)=\sum_{n\in W}f(n)=|W|\mu_W(0)$, and $\sum_{n\in W}f^2(n+k)=\sum_{n\in W}f^2(n)=|W|\gamma_W(0)$; hence

$$1-\frac{\mu_W(k)}{\mu_W(0)}=C\sqrt{1-\frac{\gamma_W(k)}{\gamma_W(0)}}, \tag{3}$$

where $C=[r(k)\sqrt{\gamma_W(0)}]/[\sqrt{2}\mu_W(0)]\approx$ constant locally. Due to this square-root relationship, the drop of $\mu(k)/\mu(0)$ from its peak value 1 is sharper (has steeper slope) than the drop of $\gamma(k)/\gamma(0)$ from its peak value 1. This explains why the morphological correlation gives sharper peaks than the linear, assuming some local stationarity.

3 Statistical Analysis

Our previous results can also be cast in a statistical framework. For example, given two (temporal or spatial) random processes $f(t)$ and $g(t)$, we define their morphological correlation by $\mu_{fg}(t_1,t_2)=\mathcal{E}\{\min[f(t_1),g(t_2)]\}$, where $\mathcal{E}\{\cdot\}$ denotes statistical expectation. If they are jointly stationary and $\tau=t_1-t_2$, then $\mu_{fg}(t_1,t_2)=\mu_{fg}(\tau)$. Considering the signal detection problem under the MAE matching criterion, minimizing $E_1(\tau)=\mathcal{E}\{|f(t+\tau)-g(t)|\}=\mathcal{E}\{f(t+\tau)\}+\mathcal{E}\{g(t)\}-2\mu_{fg}(\tau)$ as a function of the shift τ is equivalent to maximizing $\mu_{fg}(\tau)$ under various assumptions; examples include (1) stationarity of f and g, or (2) zero-mean processes f,g, or (3) normalization of both $E_1(\tau)$ and $\mu_{fg}(\tau)$ by the average of the means of f and g.

In what follows we focus on the specific example of calculating the morphological correlation for random spatio-temporal signals that are intensities of fully developed, coherent, Gaussian *speckle* patterns. These speckle images are interesting both intrinsically and in modeling some types of image noise. Let I_1 and I_2 be two random variables representing the intensities of such a speckle process at two points in time-spatial coordinates. As well-known, the marginal intensity disctributions are one-sided exponential densities $p_i(I_i)=(1/m_i)\exp(-I_i/m_i)$, $I_i\geq 0$, where $m_i=\mathcal{E}\{I_i\}$, $i=1,2$, are the mean intensities. The *linear correlation* coefficient of I_1,I_2 is $\gamma=(\mathcal{E}\{I_1I_2\}-m_1m_2)/m_1m_2$. The problem here is to calculate the *morphological correlation* $\mu=\mathcal{E}\{\min(I_1,I_2)\}$ and relate it to γ. In [5] it has been shown that, for any function $h(I_1,I_2)$,

$$\mathcal{E}\{h(I_1,I_2)\}=\sum_{n=0}^{\infty}\frac{1}{(n!)^2}\frac{\partial^{2n}\mathcal{E}\{h(I_1,I_2\}|_{\gamma=0}}{\partial m_1^n m_2^n}(m_1m_2\gamma)^n. \tag{4}$$

We apply this result for $h(I_1,I_2)=\min(I_1,I_2)$. Let

$$\mu_0=\mu|_{\gamma=0}=\frac{1}{m_1m_2}\int\int_0^{\infty}\min(I_1,I_2)e^{-\frac{I_1}{m_1}}e^{-\frac{I_2}{m_2}}dI_1dI_2.$$

It can be shown [more detailed proofs of the results in this paper will be available in a forthcoming long report] that

$$\mu_0=\frac{m_1m_2}{m_1+m_2}.$$

Hence

$$\mu=\mu_0+\sum_{n=1}^{\infty}\frac{1}{(n!)^2}\frac{\partial^{2n}\mu_0}{\partial m_1^n m_2^n}(m_1m_2\gamma)^n. \tag{5}$$

We can also arrive at a theoretically equivalent result by thresholding the intensities and using (1). Thus, by thresholding I_i at all levels a we produce the binary random variables $u(I_i-a)$, $i=1,2$, where $u(\cdot)$ is the unit step function. Then, if $h_a(I_1,I_2)=u(I_1-a)u(I_2-a)$, from (1) we obtain

$$\mu=\mathcal{E}\{\int_0^{\infty}h_a(I_1,I_2)da\}=\int_0^{\infty}\mathcal{E}\{u(I_1-a)u(I_2-a)\}da. \tag{6}$$

Finally, by using (4) for $h_a(I_1,I_2)$ with $\mathcal{E}\{h_a\}|_{\gamma=0}=\exp(-a/\mu_0)$, it follows from (6) that

$$\mu=\mu_0+\sum_{n=1}^{\infty}\frac{\gamma^n}{(n!)^2}\int_0^{\infty}\frac{a^2}{m_1m_2}e^{-\frac{a}{\mu_0}}L_{n-1}^1(\frac{a}{m_1})L_{n-1}^1(\frac{a}{m_2})da, \tag{7}$$

where $L_n^k(x)=x^{-k}e^x\frac{d^n}{dx^n}[x^{n+k}e^{-x}]$, $n=0,1,2,...$, are the associated Laguerre polynomials. Both (5) and (7) express μ as a power series expansion in terms of powers of γ and help illuminate the highly nonlinear relationship between μ and γ.

References

[1] M. Ross, H. Shaffer, A. Cohen, R. Freudberg and H. Manley, "Average Magnitude Difference Function Pitch Extractor", *IEEE Trans. ASSP-22*, Oct. 1974.

[2] E. Coyle, "Rank Order Operators and the Mean Absolute Error Criterion", *IEEE Trans. ASSP-36*, 1988.

[3] P. Maragos, "Optimal Morphological Approaches to Image Matching and Object Detection", *Proc. 2nd Int. Conf. Comp. Vis.*, Florida, Dec. 1988.

[4] J. Serra, *Image Analysis and Mathematical Morphology*, Acad. Press, 1982.

[5] H. Pedersen, "Theory of speckle-correlation measurements using nonlinear detectors", *J. Opt. Soc. Am.*, A, 1(8), Aug. 1984.