A PROBABILISTIC APPROACH TO BOUNDARY HANDLING

Anil Anthony Bharath & Jeffrey Ng Sing Kwong

Faculty of Engineering
Imperial College London
London SW7 2AZ
United Kingdom

ABSTRACT

We address the problem of boundary handling in correlation-based template matching by proposing a probabilistic model of the detection process. Whilst our approach bears similarities to those taken in deriving results in matched and subspace signal detection, it offers a new interpretation: that a dual correlator architecture provides a systematic way of handling general uncertainty, and, more specifically, the boundaries of data in signals. We also provide an extended model to deal more effectively with amplitude variations of target with respect to template. These improvements have immediate applications not only in classical matched signal detection, but also for template matching by correlation in digital image analysis and computer vision, where partial target occlusion at image boundaries remains a significant problem.

1. INTRODUCTION

The framework for presenting signal detection and localisation is similar to standard Bayesian formulations, but there are differences in notation that play a role in relating theory to correlator design. We start with the simple problem of detecting a known structure in noise. Consider samples of a continuous time or continuous space process, at location values denoted as $t_1, t_2, ..., t_N$. The scalar values of some observable quantity at these locations, $f(t_i)$, are denoted $f_1, f_2, ...f_N$ for convenience, but it will prove useful to recognise that f_1 represents the value of some measured quantity observed at location t_1, and that the various components are *ordered* in time or space.

The observation vector $\mathbf{f}_N = [f_1, f_2, ...f_N]$ is considered to contain a signal, S, that spans some subspace, or all, of R^N in a given observation. For example, the signal of interest may be present on four of the observation components, corresponding to f_1, f_2, f_3, f_4, or perhaps it might be located on components f_4, f_5, f_6, f_7. Furthermore, it may even be the case that the signal S, despite being of dimension M which is smaller than N, is *not* fully contained in R^N, but lies in some domain such as R_{N+K}, $M > K > 0$, and where, generally, $R^N \in R^{N+K}$. This implies that we may have only partial observations of the full signal structure. Desai ([1]) also addresses subspace signal detection, but without reference to correlator architecture. Furthermore, whilst Desai ([1]) provides intuitive and powerful arguments based on subspace geometry, our argument is essentially statistical.

1.1 Rhn

To enable a decision on whether and where a signal is present within an observation window, we may rely on the statistical description provided by the posterior density function $p(S, t|\mathbf{f}_N, \mathbf{h}_u)$. We write,

$$p(S, t|\mathbf{f}_N, \mathbf{h}_u) \quad = \quad \frac{p(\mathbf{f}_N|S, t, \mathbf{h}_u) \times p(S, t|\mathbf{h}_u)}{p(\mathbf{f}_N|\mathbf{h}_u)} \tag{1}$$

where t denotes a hypothetical position of the signal. In fact, t requires the definition of a point of origin in the message, S, which may, for example be its starting point or its centre. The term $\mathbf{h_u}$ is a set consisting of all remaining observation and signal parameters that affect the context, resolution, scale, amplitude and other observation and signal transformation hypotheses under which the data are observed. One may establish a hypothesis testing framework to distinguish between different signals, $S = s_1$ and the null signal, $S = s_0$, or between different non-null signals quite easily. For a countable set of M possible non-null signals, $s_1, s_2, ...s_M$ the basic architecture of signal cross-correlation arises from just such a framework; a set of $M + 1$ log likelihood signals is implicitly constructed as functions of time for each of the message hypotheses (including the null hypothesis).

For a given "message" or possible signal structure, Equation (1) represents a complete theoretical framework for addressing the inference problem, in the sense that possible answers to the task of signal detection and localisation are encapsulated in the construction of the left hand side of Equation (1). A flat density function is associated with a complete lack of certainty about where the signal is located, and a very sharply peaked density function illustrates certainty about various hypothetical positions of a specific signal.

In a practical detection situation, where a decision has to be made about the existence or absence of a specific signal (or message), one has several options available to arrive at a decision: peaks of high probability may be identified, corresponding to ML if a non-informative prior on the location of signals is used, or to MAP estimation in the more general case. If one is certain that the signal is present, and one's task is to locate it, then the positions of peaks in the posterior density function might be used. In general, one also has a measure of the certainty of these conclusions in the form of the relative sharpness (not just the height) of any peaks. In classical signal detection theory, one can find the first treatments on optimality of detection in [2], and the principled use of risk in this context by [3]. For multiple message (signal) hypotheses one may refer to [4] for a comprehensive treatment, and [5] for extensions.

2. AN EXAMPLE OF THE MODEL

The components of the observation vector, $f_i, i = 1..N$ are treated as jointly normally distributed with possibly different variances *conditional* on the signal location, t, and on its

This work was partly funded by the UK Research Council Grant GR/R87642/02.

class, s_m. To illustrate the nature of the model, we presume, first, a simple signal, s_1, consisting of a "bump" that may be positioned anywhere within the signal window:

$$p(f(t_i)|t, S = s_1, \mathbf{h}_u) = \frac{1}{\sigma_i\sqrt{2\pi}} e^{-(f_i-\mu(t_i|t,S=s_1,\mathbf{h}_u)^2)/2\sigma_i^2}, \quad i = 1, 2, ...N$$

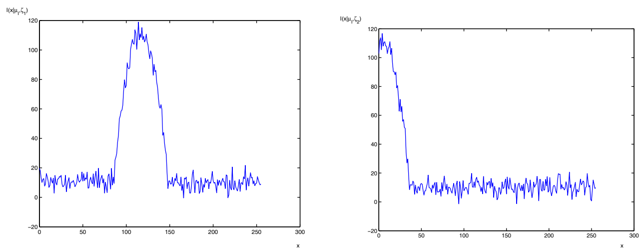
with t being the position of the peak of the bump signal, equally likely to be anywhere in the signal window,

$$t \sim \mathcal{U}[1, N] \tag{2}$$

For our "bump" signal, we select the particular form,

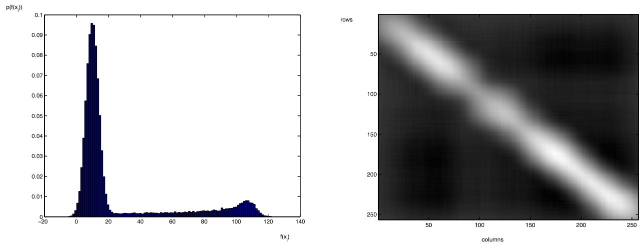
$$\mu(t_i|t, S = s_1, \mathbf{h}_u) = \begin{cases} \frac{f_{max}}{B_w^2}(t_i - t)^2 & : \quad |t_i - t| < B_w \\ 0 & : \quad \text{otherwise} \end{cases} \tag{3}$$

for some signal width $B_w > 0$. Note that with this prior on t, and this specification of the signal, partial observation or occlusion of the signal may occur. Two examples of samples drawn from such a statistical model are shown in Figure (1); these are generated by a sequential sampling technique [6] for σ_i constant. The signal template is deterministic in this simplistic model, and no variation on its shape is permitted. It is of course, perfectly possible to include other signal parameters, such as amplitude, width and so on; we shall present results of the correlation structure required for handling possible amplitude variations in Section 4.1. The choice of model is, here, the simplest that will permit new conclusions to be drawn. Figure (2) illustrates that samples drawn independently from this model have non-trivial unconditional amplitude statistics.



(a) Signal Sample 1 (b) Signal Sample 2

Fig. 1: Signals drawn from the model by a sequential sampling technique.



(a) (b)

Fig. 2: (a) Illustration of the histogram and (b) covariance structure of the signal amplitude as estimated during Monte-Carlo simulations. Note that *conditional* independence of the signal samples does not imply a diagonal (unconditional) covariance matrix.

One may now formulate the signal-detection and localisation problem, given that the "bump' signal exists, through the posterior density function,

$$p(t|f_1, f_2, ...f_N, s_1, \mathbf{h}_u) = \frac{e^{-\sum_{i=1}^{N}(f_i-\mu(t_i|t,s_1,\mathbf{h}_u))^2/2\sigma_i^2}}{(2\pi)^{N/2}\prod_{i=1}^{N}\sigma_i} \times \frac{p(t|s_1, \mathbf{h}_u)}{p(f_1, f_2, ...f_N|s_1, \mathbf{h}_u)} \tag{4}$$

An expansion of the dominant term is very instructive; denoting the log-posterior distribution by $L(t)$, setting $c_i = 1/\sigma_i^2$, and assigning all other terms to the constant K_1, we have

$$\begin{aligned} L(t) &= K_1 + \sum_{i=1}^{N} c_i f_i \mu(t_i|t, s_1, \mathbf{h}_u) \\ &- \frac{1}{2}\sum_{i=1}^{N} c_i \mu^2(t_i|t, s_1, \mathbf{h}_u) - \frac{1}{2}\sum_{i=1}^{N} c_i f_i^2 \end{aligned} \tag{5}$$

As it stands, Equation(5) does not provide a reliable detector; one needs to construct a likelihood ratio test against the possibility of other signals. Practically speaking, this is best achieved through a learning process. However, one can immediately test against the null hypothesis signal, yielding a log-likelihood posterior ratio of

$$L_r(t) = \sum_{i=1}^{N} c_i f_i \mu(t_i|t, s_1, \mathbf{h}_u) - \frac{1}{2}\sum_{i=1}^{N} c_i \mu^2(t_i|t, s_1, \mathbf{h}_u) \tag{6}$$

3. CORRELATOR ARCHITECTURE

The correlator architecture is illustrated in Figure (3). For a constant, finite and non-zero value c_i, (or $c(t_i)$), *including* for those values of t_i that correspond to locations lying outside the boundaries of the observation window, the computational structure is identical to that of Figure (III B-2) of [7]. To emphasise the difference between a multiplier that weights each data point according to its certainty, c_i, and the operation of a sliding multiplication between template and the (weighted) signal, there are some changes to the standard notation (see caption) in Figure (3). It is implicit that the number of samples of the signal f_i is N, and of the template, $\mu(t_i|t)$ is M, where in general $M < N$. For a finite observation window, it is clearly not appropriate to use the same value for c_i for $1 \leq i \leq N$ as for $i < 1$ and $i > N$; the c_i reflects one's certainty about the *data*, which tends to zero outside of the observation window.

4. SIGNIFICANCE

Equation (6) provides a correction term to template matching by standard correlation. The second term of (6) involves application of a simple non-linearity to the conditional mean function, $\mu(t_i|t, s_1, \mathbf{h}_u)$; the output of this is non-linearly "cross-correlated" against the signal specifying the certainty of the data. In standard correlator design, this certainty signal is treated as a constant [7]. However, because the variance of the signal observations tends to infinity outside of the observable area (assuming one has no prior knowledge on how the signal behaves *outside* the observable window), the result of this second cross-correlation will generally *not* be constant near to the signal boundary, but will decrease. This is illustrated by the dashed red-traces in Figures (5 (a) and (b)). The dash-dot

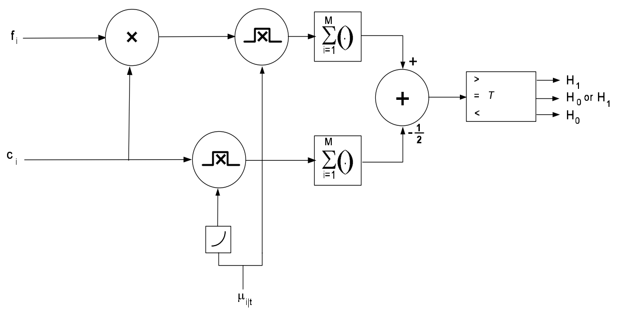


Fig. 3 : Illustrating the dual cross-correlator structure suggested by Model 1. The non-linearity applied to the template function $\mu(t_i|t)$ is a squaring operation. The multiply signal within a rectangular pulse is a sliding (windowed) multiplier. The optimum threshold, T, depends on the posterior probabilities of H_0 and H_1, and some risk function.

green trace is normal weighted cross-correlation between template and signal, for two experimental realisations of the bump signal, which is visible as a solid, noisy blue trace at the bottom of each figure. In one case, Figure (5(a)), the bump is fully visible in the signal window; in the other case, (b), the signal is partially occluded (its peak is at the right signal boundary). Note that the weighted correlation result is significantly different in the two figures, illustrating a significant correlator output drop due to the partial occlusion at the boundary; this should be compared against the behaviour of the solid, smooth, orange line, which displays a broadening, but no significant drop in amplitude. One can see that the inclusion of this second term of Equation (6) compensates effectively for the partial occlusion of the bump signal.

It should be pointed out that this follows directly from the interpretation of noise as uncertainty; outside of the signal window, one's uncertainty of the signal becomes very large. This is not equivalent to zero padding on the data channel, but rather to zero-padding on the "certainty" channel. This is a subtle, but important, distinction from the classical model of signal + noise. Figure (5) is, to be sure, a demonstration of the principle, and is not conclusive. More conclusive evidence in support of this interpretation and approach will follow in the image detection problem of Section (5).

4.1 [illegible]

The model described in the previous section does not permit amplitude scalings of the signal message to occur in an observation. In many applications of correlation, such amplitude changes are simply accepted as providing a scaling of the correlation measure. If it is desirable to have amplitude invariant measures of similarity, one approach would be to globally normalise the observation vector in some way. However, if there are other large amplitude signals, not of interest, in the observation vector, such normalisations will have the undesired effect of suppressing the correlation response to the true signal. A local normalisation is also possible which requires dividing each correlator sample by the 2-norm of the signal in a window spanning that of the template about the current signal position, but this, too, can become "badly behaved" in low-signal amplitude regions. In a related, but distinct problem, we have shown [8] how to derive an alternative correlation strategy for vector field template matching through analytic marginalisation. By adopting a similar approach to that of [8] for scalar fields, in which one specifies a Gaussian mixture model as prior on the scaling of the observations of the message relative to the template, one can derive an alternative dual-correlator structure that enhances detection performance. The full derivation is presented in [9], and the suggested architecture is illustrated in Figure (4) with a demonstration of its performance in a template matching problem in Section (5). The architecture also requires a dual correlator structure, but employs quite different non-linearities to those shown in Figure (3).

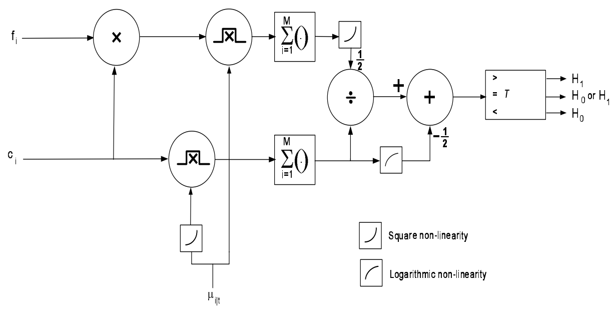
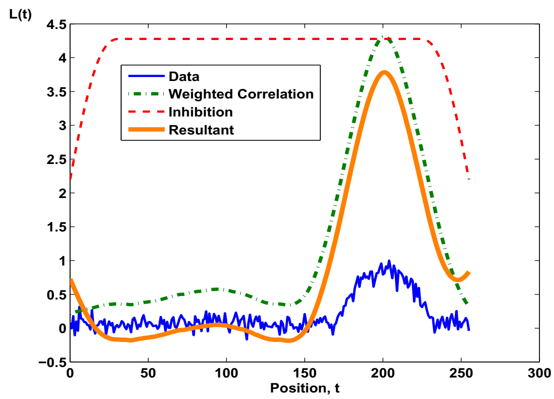


Fig. 4 : Illustrating a modified dual cross-correlator structure. This provides better performance when there is uncertainty in the relative amplitude of the target with respect to the template. This architecture also displays good boundary handling.

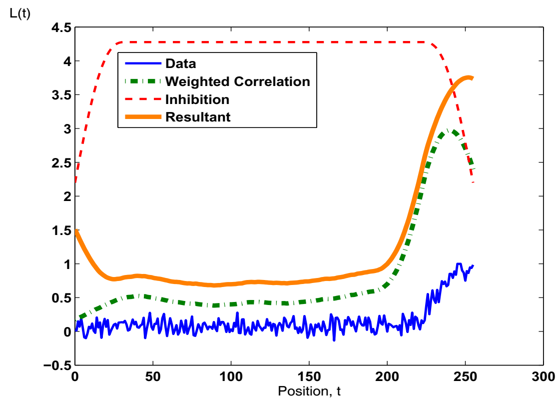
5. IMAGE TEMPLATE MATCHING

The extension of the proposed correlator design has implications for template matching in image analysis. In this section, we illustrate the improvement in detection for partially occluded structures in a simple planar template matching problem between scalar fields.

The image of Figure (6(a)) contains three coins, one of which is partially occluded. A 2D image template is constructed which consists of a circle of radius 20 pixels, with an intensity of 200. Figure (6(b)) shows the detection obtained when a thresholding is applied to the standard correlator space to best as possible capture the location of the three coins. The white '×' inlaid in red boxes show the positions of the detected peaks. Although both fully visible coins are detected, there is a splitting of the distribution peaks for one of the coins, resulting in a false detection. Whilst heuristics may be used to remove this extra peak, it represents a post-processing overhead. However, even with such post-processing, the partially occluded coin would not be detected using reasonable threshold settings. Figure (6(c)) illustrates the result of the boundary corrected cross-correlation derived in Section (2). We note that all three of the coins have been detected, but there is also a false detection due to the reflectance at the corner of the image (the background is a shiny surface). Finally, Figure (6(d)) illustrates the result of amplitude marginalised cross-correlation with boundary correction. It may be noted that all three coins are detected. In generating the results of Figures (6(b-d)), all log-likelihood posterior density estimates were renormalised to range from 0 to 1. Connected components labelling of the supra-threshold pixels was performed, followed by a simple binary morphological shrinkage operation, all of which are standard operations in image processing and have a low overhead. Thresholds were set individually for best performance in each case, but full Receiver-Operator Characteristics (ROC's) have been generated, for a different, but related problem, and are presented in ([8]).



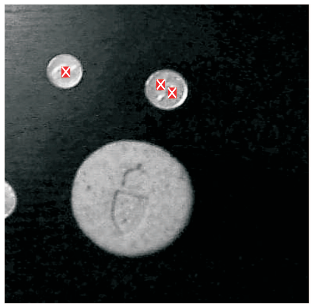
(a) Target Signal Entirely in Window



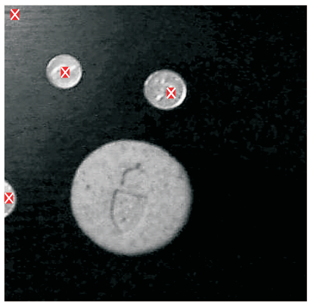
(b) Target Signal Partially Occluded

Fig. 5 : (a) (left) Cross-Correlation between a template and a conditionally independent Gaussian signal. The faint dotted line represents weighted correlation (weighting is, however, uniform within the signal window), and the red, solid line represents the inclusion of the "inhibition" terms which have an effect at the signal boundary (b)(right) The importance of correct boundary handling is illustrated by comparing the response from correlation when the signal structure ("bump") to be detected is only partially visible .

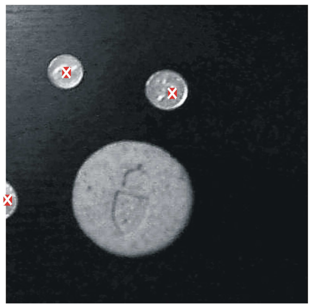
(a)



(b)



(c)



(d)

Fig. 6 : (a) Cookie and Coins Image. Note that one coin is partially occluded at the left boundary. (b) Result of cross-correlation between circular template and the image of (a). (c) Detection using boundary corrected cross-correlation. The partially occluded coin is now detected, although the reflection at the top left corner of the image leads to false detection. (d) Result of using a correlation architecture suggested by amplitude marginalisation (see text for explanation). Contrast in (d) is reduced in print because of marker at boundary which forces a white image border.

6. CONCLUSIONS

The interpretation of the signal detection problem has been addressed in many contexts. Here, we specifically deal with the problem of handling the boundaries of observation windows in an elegant fashion. We have verified that the solutions we present are plausible, and do not conflict with standard signal detection theory. The potential applications are very widespread; the problem of handling signal boundaries in object detection, particularly in image processing, is well known. We have confirmed that our suggested dual-correlator structure does indeed yield improved performance in spatial template matching. Moreover, when the underlying statistical model is extended to parameterise amplitude scalings of the template, and one marginalises over this parameter, performance may be refined even further, and this is illustrated in Section (5). We are currently extending the model to deal with conditionally independent statistics which are distinctly non-Gaussian in nature.

REFERENCES

[1] M.N. Desai and R.S. Mangoubi, "Robust Gaussian and non-Gaussian matched subspace detection," *IEEE Trans. on Signal Processing*, vol. 51, no. 12, pp. 3115–3127, Dec. 2003.

[2] W. Peterson, T. Birdsall, and W. Fox, "The theory of signal detectability," *IEEE Trans. on Inf. Theory*, vol. 4, no. 4, pp. 171–212, Sep. 1954.

[3] D. Middleton and R. Esposito, "Simultaneous optimum detection and estimation of signals in noise," *IEEE Trans. on Inf. Theory*, vol. 14, no. 3, pp. 434–444, May 1968.

[4] A. Fredriksen, D. Middleton, and V. VandeLinde, "Simultaneous signal detection and estimation under multiple hypotheses," *IEEE Trans. on Inf. Theory*, vol. 18, no. 5, pp. 607–614, Sep. 1972.

[5] C.C. Leang and D.H. Johnson, "On the asymptotics of m-hypothesis Bayesian detection," *IEEE Trans. on Inf. Theory*, vol. 43, pp. 280 – 282, Jan. 1997.

[6] Michael Evans and Tim Swartz, *Approximating Integrals via Monte Carlo and Deterministic Methods*, Oxford University Press, 2000.

[7] H. Vincent Poor, *An Introduction to Signal Detection and Estimation*, Springer-Verlag, second edition, 1994.

[8] S. Basalamah, A. Bharath, and D. McRobbie, "Contrast marginalised gradient template matching," in *Computer Vision - ECCV 2004*, 2004, vol. 3023 of *Lecture Notes In Computer Science*, pp. 417–429.

[9] Anil Anthony Bharath, "On detection by correlation in conditionally Gaussian and Gamma processes," *IEEE Trans. on Inf. Theory*, vol. Submitted, 2005.