Fast and robust detection of a known pattern in an image

Loïc Denis, André Ferrari, David Mary, Laurent Mugnier, Eric Thiébaut

Fast and robust detection of a known pattern in an image

Loïc Denis*, André Ferrari†, David Mary†, Laurent Mugnier♯, Eric Thiébaut‡

∗ Laboratoire Hubert Curien, UMR 5516 CNRS Université Jean Monnet, France
† Laboratoire Lagrange, UMR 7293 CNRS Université de Nice Sophia-Antipolis, France
♯ ONERA – The French Aerospace Lab, F-92322 Châtillon, France
‡ Observatoire de Lyon, CRAL, UMR 5574 CNRS Université Lyon 1, ENS de Lyon, France

***Abstract*—Many image processing applications require to detect a known pattern buried under noise. While maximum correlation can be implemented efficiently using fast Fourier transforms, detection criteria that are robust to the presence of outliers are typically slower by several orders of magnitude. We derive the general expression of a robust detection criterion based on the theory of locally optimal detectors. The expression of the criterion is attractive because it offers a fast implementation based on correlations. Application of this criterion to Cauchy likelihood gives good detection performance in the presence of outliers, as shown in our numerical experiments. Special attention is given to proper normalization of the criterion in order to account for truncation at the image borders and noise with a non-stationary dispersion.**

***Index Terms*—robust detection, locally most powerful test (LMP), Cauchy distribution**

I. Introduction

We consider the problem of detecting a *known pattern* from a noisy image. In its simplest formulation, this detection problem amounts to a binary hypothesis choice:

$$\begin{cases} H_0: & \boldsymbol{y} = \boldsymbol{\epsilon} \\ H_1: & \boldsymbol{y} = \alpha\, \boldsymbol{m}(\boldsymbol{x}) + \boldsymbol{\epsilon}\,, \quad \text{with } \alpha > 0 \end{cases} \tag{1}$$

where under H_0 the observation $\boldsymbol{y} \in \mathbb{R}^n$ corresponds to the realization $\boldsymbol{\epsilon}$ of a random process accounting for noise. Under hypothesis H_1, the observation is the superimposition of noise $\boldsymbol{\epsilon}$ and the pattern $\boldsymbol{m}(\boldsymbol{x}) \in \mathbb{R}^n$ of (unknown) amplitude $\alpha > 0$ and location $\boldsymbol{x}$.

In many practical cases, there are several observations y_i that display strong deviations from the model $\alpha\, m_i(\boldsymbol{x})$. These deviations may be due to non-linearities of the sensor (e.g., saturation, dead/hot pixel) or mis-modeling (e.g., structured background). The random vector $\boldsymbol{\epsilon}$ should not then be modeled by a white Gaussian distribution, but rather by a non-stationary, non-Gaussian, and perhaps correlated distribution. In this work, we keep the simplifying independence assumption but consider non-Gaussian noise distributions with heavy tails in order to derive a robust detection method.

This work has been supported by a grant from the French CNRS (DETECTION, DEFI IMAG'In 2015)

When the measurement and modeling errors $\boldsymbol{\epsilon}$ are centered and can be considered Gaussian, with a known covariance matrix $\boldsymbol{\Gamma}$, the generalized likelihood ratio test takes the form of a matched filter [1]:

$$T_{\text{Corr}}(\boldsymbol{x}) = \max\left(\frac{\boldsymbol{m}(\boldsymbol{x})^{\text{t}}\boldsymbol{\Gamma}^{-1}\boldsymbol{y}}{\sqrt{\boldsymbol{m}(\boldsymbol{x})^{\text{t}}\boldsymbol{\Gamma}^{-1}\boldsymbol{m}(\boldsymbol{x})}}, 0\right) \underset{H_0}{\overset{H_1}{\gtrless}} \gamma\,, \tag{2}$$

when taking into account the prior information that $\alpha \geq 0$. Evaluating T_{Corr} for different locations $\boldsymbol{x}$ gives a correlation map. Correlation-based methods offer optimal performance under Gaussian noise [1]. It is well known, however, that the performance strongly deteriorates in the presence of large deviations (outliers), see [2] and figure 1.

Robust image matching is a widely studied topic. A common approach starts by detecting feature points, then computes local descriptors such as scale-invariant feature transforms (SIFTs) [3] before matching pairs of descriptors. Robust matching is achieved by restricting the matching to feature points, and by using well-chosen (i.e., invariant) descriptors [4].

In order to detect patterns even at low signal-to-noise (SNR) regimes, a "pixel-based" approach is often better suited. The fast correlation method proposed in [2] for image registration replaces the quadratic penalty typical of maximum likelihood under a Gaussian assumption by a bounded penalty of the form $\psi(\Delta) = 1 - \cos(\Delta)$. The reason for this specific choice of penality is that it leads to a fast implementation based on FFTs.

In a signal processing context, locally optimal detectors (i.e., detectors statistically optimal for weak signals) have been developed for several noise distributions, see the seminal papers by Capon [5] and by Miller and Thomas [6]. These detectors have been applied recently to the detection of marks in image watermarking [7], [8]. In this paper, we derive the general expression of locally optimal detectors for problem (1), including the normalization of the test statistics required for the calibration of the threshold. This normalization is crucial to account for truncation at the image borders and prevent the introduction of bias for objects centered close to or beyond the image borders. It is also necessary when considering noise with a non-stationary dispersion. We show that the expression

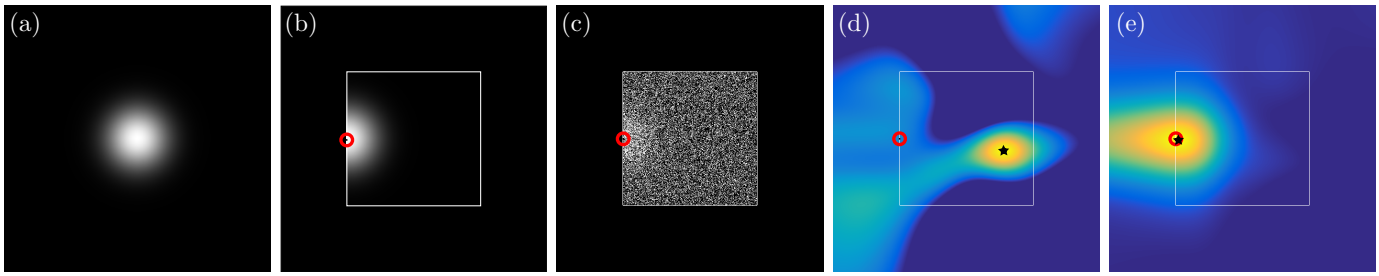

Fig. 1. Detection of a pattern in the presence of strong outliers: correlation vs. robust correlation. (a) the reference pattern; (b) noiseless signal: measurements are available only inside the white square, the center of the actual pattern is marked with a red circle; (c) noisy data (SNR=1 for 95% of the pixels, and SNR=0.01 for the remaining 5%) shown using the same display range as image (b); (d) normalized correlation map T_{Corr} (maximum marked with a star); (e) proposed robust-correlation map.

of the criterion offers a fast implementation based on pixel-wise operations and discrete correlations. We give its close-form expression under Cauchy noise and prove its Gaussian approximation in the asymptotic regime. Finally, we illustrate its performance on a numerical experiment.

II. Proposed robust detector

A. Motivation

Several tests have been developed in the framework of detection theory. The three most well known are the (generalized) likelihood ratio, Wald's statistic and Rao's score, see [9], [10]. Robust detection criteria can readily be derived from these tests by modeling the distribution of errors $\mathrm{p}(\boldsymbol{\epsilon})$ with a heavy-tailed distribution. However, if the location $\boldsymbol{x}$ of the pattern is not known (at least approximately) beforehand, robust detection criteria are generally computationally very demanding because they require estimating the amplitude α during the exhaustive search for location $\boldsymbol{x}$. Maximum likelihood estimation of α under non-Gaussian distributions (an M-estimator) typically requires resorting to an iterative procedure [11]. Testing all pixel locations thus leads to prohibitive computation times when the pattern $\boldsymbol{m}$ has a large spatial extension.

A very interesting feature of Rao's score is that it does not require the estimation of the amplitude of the pattern. When the (unknown) amplitude of the pattern to detect is always non-negative, Rao's score corresponds to the locally most powerful (LMP) detector, i.e., it is asymptotically optimal for weak signals (when α tends to zero). The LMP test is given by [1]:

$$T_{\text{LMP}}(\boldsymbol{y}) = \left.\frac{\partial \log \mathrm{p}(\boldsymbol{y}|\alpha, H_1)}{\partial \alpha}\right|_{\alpha=0} \cdot I_0^{-1/2} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma\,, \tag{3}$$

where I_0 denotes Fisher information on parameter α, evaluated at $\alpha = 0$:

$$I_0 = \mathbb{E}\left[-\frac{\partial^2 \log \mathrm{p}(\boldsymbol{y}|\alpha, H_1)}{\partial \alpha^2}\right]\Bigg|_{\alpha=0}\,. \tag{4}$$

The LMP test does not require the estimation of the amplitude α, but only the evaluation of the slope of the log likelihood at $\alpha = 0$ (i.e., under H_0) and the computation of the expectation of the curvature of the log likelihood for the Fisher information. In the case of Gaussian noise, the LMP detector leads to the same expression (2) as the generalized likelihood ratio. We investigate the use of this test with heavy-tailed distributions as a detection criterion robust to the presence of outliers.

The general expression of LMP detectors has been derived by Miller and Thomas [6], without the normalization factor I_0. It is worth to note that, as already mentioned in the introduction, the normalizing factor I_0 plays a crucial part in image processing. I_0 is a function of $\boldsymbol{m}(\boldsymbol{x})$ which generally depends on the position $\boldsymbol{x}$. While the threshold γ can be calibrated once and for all, using Monte-Carlo simulations or asymptotic distribution of the test statistic, in the center of the image, i.e. where $\boldsymbol{m}$ is constant, a different calibration of γ will be required for each position of the pattern on the border of the image. Moreover, if the noise displays a non-stationary dispersion, the normalization I_0 also varies spatially.

B. Derivation of the detector

We define the scaled residuals $\boldsymbol{t}(\alpha, \boldsymbol{x})$ by

$$\forall i,\ t_i(\alpha, \boldsymbol{x}) = \frac{y_i - \alpha\, m_i(\boldsymbol{x})}{s_i} \tag{5}$$

where we assume that the scaling factors $s_i > 0$ can be chosen such that under H_0, random variables $t_i(0, \boldsymbol{x})$ are independent and identically distributed (i.i.d.). We thus consider a slightly more general case than that of i.i.d. noise. By considering non-constant scaling factors s_i, one can account for noise with non-stationary dispersion.

We further assume that the likelihood of the data can be written under the form

$$\mathrm{p}(\boldsymbol{y}|\alpha, \boldsymbol{m}(\boldsymbol{x})) \propto \textstyle\prod_i \exp(-\varphi(t_i))\,, \tag{6}$$

with φ twice differentiable.

To compute the first factor of the test (3), we begin by computing the partial derivative of the log-likelihood with respect to the amplitude α of the pattern:

$$\frac{\partial \log \mathrm{p}(\boldsymbol{y}|\alpha, H_1)}{\partial \alpha} = \sum_i \frac{m_i(\boldsymbol{x})}{s_i} \varphi'(t_i)\,, \quad \text{with } \varphi' = \frac{\partial \varphi}{\partial t} \tag{7}$$

which gives:

$$\left.\frac{\partial \log \mathrm{p}(\boldsymbol{y}|\alpha, H_1)}{\partial \alpha}\right|_{\alpha=0} = \sum_i \frac{m_i(\boldsymbol{x})}{s_i} \varphi'(y_i/s_i)\,. \tag{8}$$

The second part requires the computation of Fisher information. We therefore compute the second order partial derivative:

$$\frac{\partial^2 \log \mathrm{p}(\boldsymbol{y}|\alpha, H_1)}{\partial \alpha^2} = -\sum_i \frac{m_i^2(\boldsymbol{x})}{s_i^2} \varphi''(t_i)\,. \tag{9}$$

Fisher information on parameter α, for $\alpha = 0$, is given by:

$$I_0(\boldsymbol{x}) = \sum_i \frac{m_i^2(\boldsymbol{x})}{s_i^2} \mathbb{E}[\varphi''(y_i/s_i)]\,. \tag{10}$$

Since under our initial hypothesis $t_i(0) = y_i/s_i$ are i.i.d., the expectation is constant:

$$\forall i,\ \mathbb{E}[\varphi''(y_i/s_i)] = \underbrace{\int_{-\infty}^{\infty} \varphi''(t) \exp(-\varphi(t))\,\mathrm{d}t}_{\beta}\,, \tag{11}$$

which gives the final expression of the LMP test:

$$T_{\mathrm{LMP}}(\boldsymbol{x}) = \tau(\boldsymbol{x}) \sum_{i=1}^{n} \frac{m_i(\boldsymbol{x})}{s_i} \varphi'(y_i/s_i)\,, \tag{12}$$

$$\text{with} \qquad \tau(\boldsymbol{x}) = \left(\beta \sum_{i=1}^{n} \frac{m_i^2(\boldsymbol{x})}{s_i^2}\right)^{-1/2}. \tag{13}$$

We have written explicitly the dependency on the location $\boldsymbol{x}$ of the model in order to emphasize that the normalization coefficient $\tau(\boldsymbol{x})$ varies in the field of view, either because the model is shifted with respect to non-constant scaling factors s_i, or because truncation and/or deformation of the model occur. We further discuss the practical implication of the spatial variation of $\tau(\boldsymbol{x})$ in paragraph II-E.

C. Asymptotic normality under H_0

The model under consideration is not i.i.d. and consequently asymptotic normality of the LMP test statistic under H_0 cannot be assumed. The expression of the test obtained in equation (12) can be re-written under the form:

$$T_{\mathrm{LMP}}(\boldsymbol{x}) = \frac{1}{\sqrt{\sum_{i=1}^n c_i^2(\boldsymbol{x})}} \sum_{i=1}^{n} c_i(\boldsymbol{x})\, u_i\,, \tag{14}$$

$$\text{with}\ \ c_i(\boldsymbol{x}) = \frac{m_i(\boldsymbol{x})}{s_i} \ \ \text{and} \ \ u_i = \varphi'(y_i/s_i)/\sqrt{\beta}\,. \tag{15}$$

The expectation of random variables u_i under H_0 is 0, while their variance is equal to 1. By application of lemma 11.3.3 from [10], a sufficient condition for the test statistic $T_{\mathrm{LMP}}(\boldsymbol{x})$ to converge to a standard normal distribution under H_0 is that:

$$\lim_{n\to\infty} \frac{\sum_{i=1}^n c_i^2(\boldsymbol{x})}{\max_{i=1..n} c_i^2(\boldsymbol{x})} = \infty \tag{16}$$

which amounts to requiring that additional observations increase Fisher information so that it is unbounded.

A threshold can thus be set to reach a prescribed false alarm rate based on a normal approximation of the sample distribution of $T_{\mathrm{LMP}}(\boldsymbol{x})$.

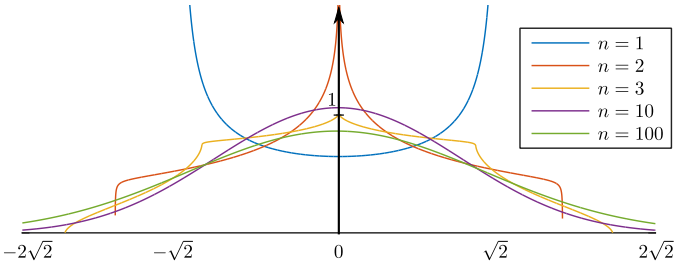

Fig. 2. Asymptotic normality of the test under Cauchy noise: probability distribution of T_{LMP} for different values of n, under H_0, with $s_i = 1$ and $m_i = 1$ for all i, and $\alpha = 1$.

D. Derivation under Cauchy noise

Cauchy distribution has heavy tails that can account for the presence of strong outliers in the data. The negative log-likelihood for a standard Cauchy distribution is given by:

$$\varphi(t) = \log(t^2 + 1) + \log(\pi)\,, \tag{17}$$

and its derivative is

$$\varphi'(t) = \frac{2t}{t^2 + 1}\,. \tag{18}$$

By application of equation (12), we obtain the expression of the robust detection test:

$$T_{\mathrm{LMP}}(\boldsymbol{x}) = 2\tau(\boldsymbol{x}) \sum_{i=1}^{n} \frac{m_i(\boldsymbol{x})}{s_i} \frac{y_i/s_i}{(y_i/s_i)^2 + 1}\,, \tag{19}$$

where the constant β in the normalization factor $\tau(\boldsymbol{x})$ can be computed in closed form [8]: $\beta = 1/2$.

Note that since the derivative of the log-likelihood of Cauchy distribution is bounded by ± 1, the random variables $u_i = \varphi'(y_i/s_i)/\sqrt{\beta}$ are uniformly bounded (both under H_0 and H_1). Moreover, the following expressions for the mean and the variance of u_i can be derived:

$$\mathbb{E}[u_i] = \frac{2\sqrt{2}\alpha s_i m_i(\boldsymbol{x})}{4s_i^2 + \alpha^2 m_i^2(\boldsymbol{x})}\,,\ \mathrm{Var}[u_i] = \frac{4s_i^2}{4s_i^2 + \alpha^2 m_i^2(\boldsymbol{x})} \tag{20}$$

and the probability density function of u_i is:

$$\mathrm{p}(u_i) = \begin{cases} \frac{\sqrt{2}}{2\pi}\left[1 - 4\left(\frac{u_i}{2\sqrt{2}} - \frac{\alpha m_i(\boldsymbol{x})}{s_i}\right)^2\right]^{-\frac{1}{2}} & \text{if } u_i \in]-\sqrt{2}, \sqrt{2}[\\ 0 & \text{otherwise.} \end{cases} \tag{21}$$

Based on (21), we can represent the distribution of the test statistic T_{LMP} for several values of n. Figure 2 illustrates the convergence to the standard normal distribution under H_0.

Asymptotic normality of T_{LMP} under H_0 can also be proved in the Cauchy noise case as a direct consequence of Lindeberg's theorem: every uniformly bounded sequence of mutually independent random variables obeys the central limit theorem, provided that the variance of the sum be unbounded (see [12]). Application of this result to the variables $m_i(\boldsymbol{x})u_i/s_i$ in (19) proves asymptotic normality of (19) when $\sum_{i=1}^n c_i^2(\boldsymbol{x}) = \infty$.

Note that when $\alpha > 0$, if $c_i(\boldsymbol{x})$ is upper bounded, the variance (20) is lower bounded. Hence when $\sum_{i=1}^n c_i^2(\boldsymbol{x}) = \infty$ the sum of variances of the variables $m_i(\boldsymbol{x})u_i/s_i$ in (19)

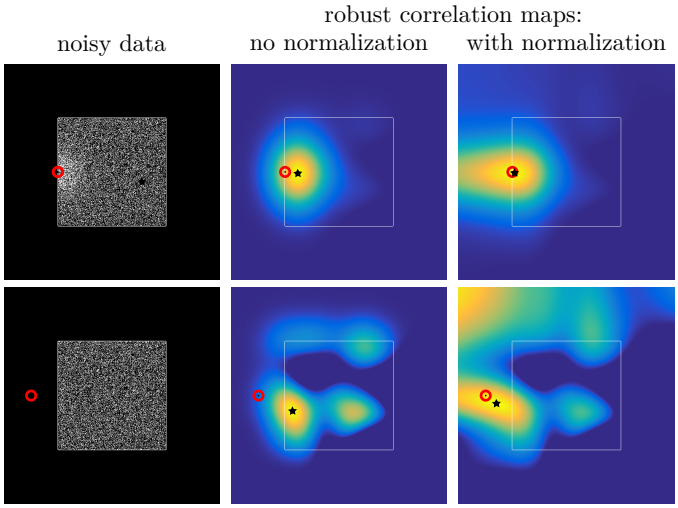

Fig. 3. Importance of proper normalization of the detection criterion: detection of a Gaussian pattern centered on the red circle, without normalization (second column) or with (last column) normalization.

is also unbounded under H_1. This proves that the sum of the $m_i(\boldsymbol{x})u_i/s_i$, properly normalized, is also asymptotically normal under H_1. This justifies a Gaussian approximation of the T_{LMP} distribution under H_1 in the finite samples case. The asymptotic probability of detection and probability of false alarm can thus be computed based on the Gaussian cumulative distribution function.

E. Map of the robust-detection criterion

The expression obtained in equations (12) and (13) follows the general form of locally optimal (LO) detectors [5], [6]:

$$T_{\text{LMP}}(\boldsymbol{x}) = \tau(\boldsymbol{x}) \cdot \boldsymbol{m}(\boldsymbol{x})^{\text{t}}\, \boldsymbol{g}(\boldsymbol{y})\,, \tag{22}$$

with $\boldsymbol{g}$ a non-linear but separable transformation applied on the vector of observed amplitudes $\boldsymbol{y}$ before correlating with the model $\boldsymbol{m}$:

$$\forall i,\ [\boldsymbol{g}(\boldsymbol{y})]_i = \frac{1}{s_i}\varphi'(y_i/s_i)\,. \tag{23}$$

and in the special case of Cauchy noise, $[\boldsymbol{g}(\boldsymbol{y})]_i = 2\,y_i/(y_i^2 + s_i^2)$. Input values y_i that are small with respect to s_i are left unchanged (up to a constant scaling factor), while larger values of y_i are attenuated in order to reduce the influence of outliers.

An attractive feature of the locally optimal tests for image processing is that they offer fast implementations based on pixel-wise operations and discrete correlations that can be computed using FFTs:

$$T_{\text{LMP}}(\boldsymbol{x}) = \frac{\boldsymbol{m} \circledast (\boldsymbol{g}(\boldsymbol{y}) \odot \boldsymbol{w})}{\sqrt{\beta}\sqrt{(\boldsymbol{m}^2) \circledast (\boldsymbol{w}/\boldsymbol{s}^2)}} \tag{24}$$

where $\boldsymbol{w}$ is a binary mask indicating the region where pixels are measured, $\boldsymbol{s}$ is a vector containing the scale factors s_i, $\odot$ denotes the element-wise product, $\circledast$ denotes the bi-dimensional correlation, and the divisions as well as the square root and squares are applied element-wise. Note that the non-linear transform $\boldsymbol{g}$ is fast to compute since it is separable.

Figure 3 illustrates the importance of the normalization (i.e., denominator in equation (24)) when computing maps of the robust-detection criterion. Without proper normalization, the maximum of the criterion is always reached inside the field of view, which biases the detection.

F. Estimation of the s parameter in the case of Cauchy noise

The application of the robust criterion derived for Cauchy distribution requires the estimation of scaling factors s_i. We consider in this paragraph the estimation of a constant parameter s. When the scaling factors s_i are chosen constant and equal to s, the expression of the robust detector under Cauchy noise simplifies into:

$$T_{\text{LMP}}(\boldsymbol{x}) = \frac{2\sqrt{2}}{\|\boldsymbol{m}(\boldsymbol{x})\|_2} \sum_i m_i(\boldsymbol{x}) \frac{y_i/s}{(y_i/s)^2 + 1}\,. \tag{25}$$

In the context of the detection of weak signals, we suggest estimating s from the data by considering that noise is dominating, or that data where the pattern is not present are available. Maximum likelihood estimation of s is done by finding the unique root of the derivative of the log-likelihood, i.e., by solving the non-linear equation [13]:

$$\sum_i \frac{2\,s^2}{s^2 + y_i^2} = n \tag{26}$$

where n is the dimension of the observation vector $\boldsymbol{y}$. Equation (26) can be solved efficiently by bisection, starting from the initial value for s: $\sum_i |y_i|/n$, or by using Brent's method [14]. This estimation algorithm is preferred to algorithms relying on order statistics, such as [15], for its computational efficiency.

III. Computer simulations

We illustrate the detection performance of the proposed criterion on a numerical simulation. The reference pattern is Gaussian (fig. 1(a)). The observations $\boldsymbol{y}$ are corrupted by Gaussian mixture noise: 95% of the pixels noise is Gaussian with an SNR of 1 and the remaining 5% is Gaussian with an SNR of 0.01 (outliers). We illustrate in figure 1(e) that the map of the detection criterion values defined in equation (24) reaches a maximum close to the true location of the pattern, while standard correlation, i.e., the detector defined in equation (2), gives a maximum at an incorrect location. As illustrated in figure 3, proper normalization is crucial to account for possible truncation of the pattern. If the normalization is missing, a bias in the localization appears.

The detection performance is further compared by computing the receiver operating characteristic (ROC) curve from 100000 random realizations in the same setting, figure 4. We compare the performance of several detectors: (i) the Neyman-Pearson idealized detector that uses the true distribution of noise (a mixture of two Gaussians) and the knowledge of the exact value of the amplitude α; (ii) the Generalized Likelihood Ratio (GLR) for a Cauchy distribution (parameter s is estimated from the data, even when $\alpha \neq 0$, using the method described in paragraph II-F); (iii) the proposed robust detector, using the same estimation of Cauchy distribution parameter s; (iv) the proposed robust detector with the best possible tuning

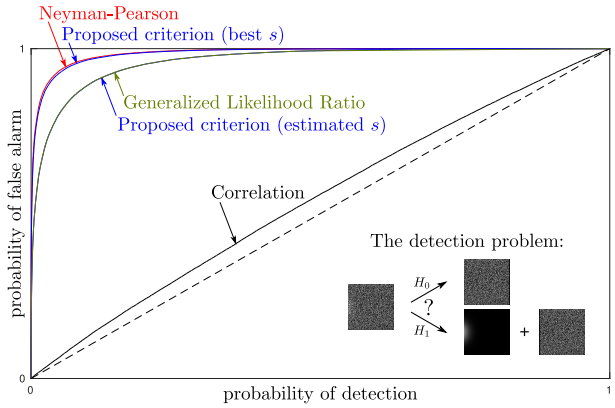

Fig. 4. ROC curves for several detectors. The curves of the GLR and the proposed LMP test are almost perfectly superimposed.

TABLE I
LOCALIZATION ACCURACY OF SEVERAL CRITERIA

Criterion	**Localization accuracy** (standard deviation in pixels)
Neyman-Pearson	2.2
Correlation	64.5
Generalized Likelihood Ratio (Cauchy)	2.8
Proposed criterion (Cauchy)	2.9

of s (i.e., which maximizes the area under the ROC curve); and (v) the usual correlation criterion given in equation (2).

From the ROC curves, it can be observed that correlation performs poorly because of the presence of outliers. Robust criteria behave much better. The GLR and the proposed LMP criteria offer very similar performance, the latter being much easier to implement since it does not require to estimate α. Note that the distribution of noise in the numerical experiment (Gaussian mixture) differs from the hypothesized distribution in the derivation of the test, without significantly impacting the performance: with the best possible choice of s, the performance of our criterion almost reaches that of Neyman-Pearson's detector.

Beyond correct detection, it is generally necessary in applications to both *detect* and *locate* the pattern. The same noise model as in figure 4 is considered, only with a stronger SNR (95% of the pixels with an SNR of 8 and the remaining 5% with an SNR of 0.08). The pattern corresponds to the truncated Gaussian with a standard deviation equal to 40 pixels that is depicted in figure 4. The location of the pattern along the y-axis is estimated by finding (by exhaustive search) the location that maximizes each test statistic. The x-axis location is kept constant and equal to the correct location. The standard deviation of the location, computed over 1000 random trials, is reported in table I for each criterion. In all cases, we found the bias to be negligible compared to the standard deviation. The Neyman-Pearson criterion, based on the knowledge of the exact distribution of noise and of the true amplitude α of the pattern gives the best accuracy. Maximizing this criterion amounts to estimating the y location with the maximum likelihood estimator using the exact likelihood and knowledge of the amplitude α. Maximizing the normalized correlation, i.e., maximum likelihood estimation under the hypothesis of a stationary Gaussian noise, gives a very strong localization error (the error is multiplied by 29). Such a non-robust criterion is not adapted to detection and estimation problems in the presence of strong outliers. Maximizing the GLR under the hypothesis of a Cauchy distribution, with the parameter s of Cauchy distribution estimated from the data, slightly degrades the accuracy compared to the (idealized) Neyman-Pearson criterion. This estimation method corresponds to the joint maximum likelihood estimation of the location y and amplitude α under an hypothesized Cauchy distribution, i.e., an M-estimator. Finally, maximizing the proposed criterion (the LMP) gives a comparable degradation of the accuracy. The huge difference in practice is that maximization of the LMP statistic does not require to jointly estimate α. It can thus be performed very efficiently on a pixel grid by FFTs.

IV. CONCLUSION

We derived the general expression of a robust detection criterion based on the theory of locally optimal detectors. We have shown the importance of the normalization term for image processing applications and have given a fast algorithm to compute detection maps. Application of the criterion to Cauchy distribution leads to a simple expression. The scale parameter of Cauchy distribution can be estimated directly from the data and the obtained detector offers a significant improvement compared to classical correlation while retaining its computational efficiency.

REFERENCES

[1] S Kay, "Fundamentals of statistical signal processing, Vol. II: Detection Theory," *Prentice Hall*, 1998.

[2] A Fitch, A Kadyrov, W Christmas, and J Kittler, "Fast robust correlation," *IEEE Trans. Image Process.*, vol. 14, no. 8, 2005.

[3] David G Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[4] J-M Morel and G Yu, "ASIFT: A new framework for fully affine invariant image comparison," *SIAM J. Imaging Sciences*, vol. 2, no. 2, pp. 438–469, 2009.

[5] Jack Capon, "On the asmptotic efficiency of locally optimum detectors," *IRE Trans. Inf. Theory*, vol. 7, no. 2, pp. 67–71, 1961.

[6] J. Miller and John B. Thomas, "Detectors for discrete-time signals in non-Gaussian noise," *IEEE Trans. Inf. Theory*, vol. 18, no. 2, 1972.

[7] Y. Bian and S. Liang, "Locally optimal detection of image watermarks in the wavelet domain using Bessel K form distribution," *IEEE Trans. Image Process.*, vol. 22, no. 6, pp. 2372–2384, 2013.

[8] A. Briassouli and M.G. Strintzis, "Locally optimum nonlinearities for DCT watermark detection," *IEEE Trans. Image Process.*, vol. 13, no. 12, pp. 1604–1617, 2004.

[9] CR Rao, "Score test: historical review and recent developments," in *Advances in Ranking and Selection, Multiple Comparisons, and Reliability*, pp. 3–20. Springer, 2005.

[10] Erich L Lehmann and Joseph P Romano, *Testing Statistical Hypotheses*, Springer, 2010.

[11] Peter J Huber and Elvezio M Ronchetti, *Robust statistics; 2nd ed.*, Wiley Series in Probability and Statistics. Wiley, Hoboken, NJ, 2009.

[12] Willliam Feller, *An introduction to probability theory and its applications*, vol. 2, John Wiley & Sons, 1971.

[13] Gerald Haas, Lee Bain, and Charles Antle, "Inferences for the Cauchy distribution based on maximum likelihood estimators," *Biometrika*, vol. 57, no. 2, pp. 403–408, 1970.

[14] Richard P Brent, *Algorithms for minimization without derivatives*, Courier Corporation, 2013.

[15] O. Y. Kravchuk and P. K. Pollett, "Hodges-Lehmann scale estimator for Cauchy distribution," *Communications in Statistics - Theory and Methods*, vol. 41, no. 20, pp. 3621–3632, 2012.