Object detection in a cluttered background with a noisy image

P.M. Aguilar-González[1], V. Kober[1,2], I.A. Ovseyevich[2]

[1] *Department of Computer Science, Division of Applied Physics CICESE, Ensenada, Mexico*

[2] *Institute for Information Transmissions Problems of RAS Moscow, Bolshoi Karetnii 19*

e-mail: `vkober@cicese.mx,vitally@iitp.ru`

Abstract

In this paper we present a correlation filter designed with a single noisy reference image in which the target is placed at unknown coordinates and its shape is unknown. The filter is designed by optimizing the peak-to-output energy criterion for a signal model that considers the presence of additive noise in the reference image and a disjoint cluttered background in the input scene. The performance of the proposed filter is compared with that of common filters by means of computer simulations.

1 Introduction

The problems of detection and location estimation with correlation filters can be solved by first searching the highest peaks in the correlation plane and then by using the coordinates of the peaks as estimates of the target location [3]. The nonoverlapping signal model has been used to represent objects located in cluttered backgrounds [2]. Common correlation filters are designed by maximizing different performance criteria. However, the resulting filters require explicit knowledge of all model parameters. Recently, correlation filters were proposed for a signal model that takes into account the presence of additive noise both in the reference image and in the input scene [1]. In this work, we propose a correlation filter derived from a model in which additive noise degrades the reference image, and the input scene contains a nonoverlapping cluttered background. This model can represent situations in which a target is taken in a controlled environment, thus no clutters are present in the reference image. The presentation is organized as follows. In section 2 we introduce a signal model and derive an optimum filter. In section 3 we present and discuss computer simulation results, as well as estimations required for the filter implementation. Finally, section 4 summarizes our conclusions.

2 Filter design

A signal model that considers additive noise in the reference image and a disjoint noisy cluttered background of the input scene is given by

$$r(x) = t(x - x_r) + n_1(x), \quad (1)$$

$$s(x) = t(x - x_s) + \bar{w}(x - x_s)\, c(x) + n_2(x), \quad (2)$$

where $r(x)$ and $s(x)$ represent the reference image and the input scene, respectively; x_r and x_s are the coordinates of the target in the reference image and input scene, respectively; $\bar{w}$ denotes the inverse region of support for the target, that is, the target shape; $c(x)$ is the cluttered background, and $n_1(x)$ and $n_2(x)$ are additive noise in the reference image and input scene, respectively. We use the following assumptions: $n_1(x)$, $n_2(x)$ and $c(x)$ are considered as realizations of stationary random processes with known first and second order statistics; x_r and x_s are random variables; all variables and stochastic processes are statistically independent of each other; the desired filter is of the form $H(\omega) = A(\omega) R(\omega)$ in the frequency domain.

The peak-to-output energy ratio (POE) criterion is used for filter design. The POE [2] is formally defined as

$$\mathrm{POE} = \left|\mathrm{E}\left\{y(x_0)\right\}\right|^2 / \mathrm{E}\left\{\overline{|y(x)|^2}\right\}, \tag{3}$$

where $\mathrm{E}\{\cdot\}$ denotes statistical averaging, $y(x)$ is the correlation plane, x_0 is the coordinate of the output peak, and the overbar denotes spatial averaging. The derived filter maximizes this criterion and tends to produce sharp peaks at the target location while the background is rejected. The location x_0 is expected to be $x_s - x_r$, which is in the close vicinity of the true location of the target as long as the target is reasonably centered in the reference image, that is, $x_r \approx 0$. The expected output peak and output energy can be calculated as follows:

$$\mathrm{E}\left\{y(x_0)\right\} = \mathrm{E}\left\{\frac{1}{2\pi}\int_{-\infty}^{\infty} H(\omega) S(\omega) \exp(i\omega x_0)\, d\omega\right\}, \tag{4}$$

$$\mathrm{E}\left\{\overline{|y(x)|^2}\right\} = \mathrm{E}\left\{\frac{1}{\mathcal{L}}\int_{-\infty}^{\infty} |y(x)|^2\, dx\right\} = \mathrm{E}\left\{\frac{1}{2\pi\mathcal{L}}\int_{-\infty}^{\infty} |H(\omega) S(\omega)|^2\, d\omega\right\}, \tag{5}$$

where $\mathcal{L}$ denotes the spatial extent of the signal, and eq. 5 holds owing to the Parseval's theorem. Because of the linearity of the expectation operator and the statistical independence of the reference image and the input scene, we can further rewrite

$$\mathrm{E}\left\{y(x_0)\right\} = \frac{1}{2\pi}\int_{-\infty}^{\infty} A(\omega)\, \mathrm{E}\left\{R(\omega)\right\} \mathrm{E}\left\{S(\omega)\right\} d\omega, \tag{6}$$

$$\mathrm{E}\left\{\overline{|y(x)|^2}\right\} = \frac{1}{2\pi\mathcal{L}}\int_{-\infty}^{\infty} |A(\omega)|^2\, \mathrm{E}\left\{|R(\omega)|^2\right\} \mathrm{E}\left\{|S(\omega)|^2\right\} d\omega. \tag{7}$$

By substituting eqs. 6 and 7 into eq. 3, we get

$$\mathrm{POE} = \frac{\left|\frac{1}{2\pi}\int_{-\infty}^{\infty} A(\omega)\, \mathrm{E}\left\{R(\omega)\right\} \mathrm{E}\left\{S(\omega)\right\} d\omega\right|^2}{\frac{1}{2\pi\mathcal{L}}\int_{-\infty}^{\infty} |A(\omega)|^2\, \mathrm{E}\left\{|R(\omega)|^2\right\} \mathrm{E}\left\{|S(\omega)|^2\right\} d\omega}. \tag{8}$$

By applying the Cauchy-Schwartz inequality, the value of $A(\omega)$ that maximizes the POE is given by

$$A^*(\omega) = \frac{\alpha \mathrm{E}\left\{R(\omega)\right\} \mathrm{E}\left\{S(\omega)\right\}}{\mathrm{E}\left\{|R(\omega)|^2\right\} \mathrm{E}\left\{|S(\omega)|^2\right\}}, \tag{9}$$

where α is a normalization constant and * denotes complex conjugate. Finally, using eqs. 1 and 2 we obtain the transfer function for the filter that maximizes the POE:

$$H^*(\omega) = \frac{\alpha R^*(\omega) T_1(\omega) T_2(\omega)}{\left[|T_1(\omega)|^2 + N_1(\omega)\right]\left[|T_2(\omega)|^2 + (2\pi)^{-1} C^0(\omega) * \left|\bar{W}(\omega)\right|^2 + N_2(\omega)\right]}, \quad (10)$$

where $T_1(\omega) = T(\omega) + \mu_1 \delta(\omega)$ and $T_2(\omega) = T(\omega) + \mu_c \bar{W}(\omega) + \mu_2 \delta(\omega)$, are the expected values of the reference image and the input scene, respectively, $C^0(\omega)$, $N_1(\omega)$ and $N_2(\omega)$ are the spectral densities of the cluttered background, reference image noise and input scene noise, respectively. If the reference image is noise free, the resulting filter is the Generalized Optimum Filter (GOF) [2]. We shall refer our filter as $\mathrm{GOF_{AN}}$ as it is a generalization of the GOF filter when the input scene is corrupted by additive noise, and we have a nonoverlapping input scene.

3 Computer simulations

In this section we present computer simulation results. The performance of the proposed filters is evaluated in terms of discrimination capability (DC) defined as

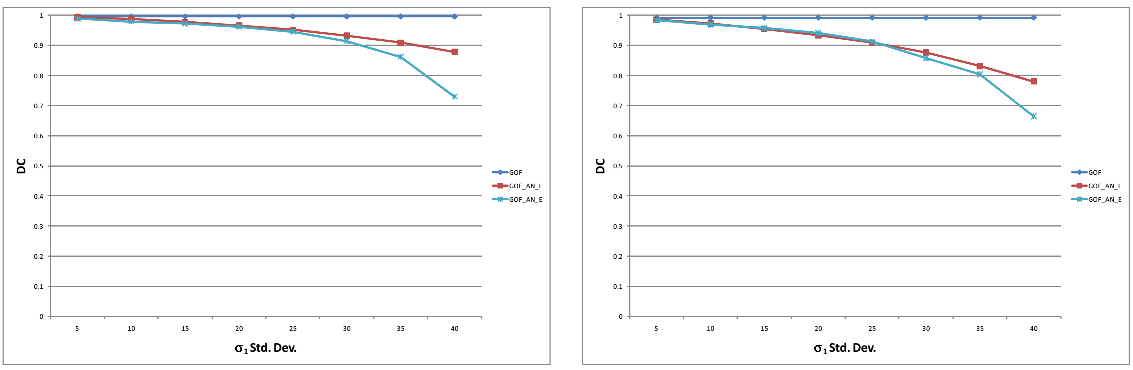
$$\mathrm{DC} = 1 - \left|y^c_{max}\right|^2 / \left|y^t_{max}\right|^2, \quad (11)$$

where y^c_{max} and y^t_{max} represent the maxima of the correlation output in the background area and in the target area, respectively. Ideally, the DC of unity is desired, while a negative DC indicates a failure to recognize the target. Images used in the experiments have size 256×256 and intensity values in the range [0-255]. We use the images shown in Figs. 1a and 1b as target and background, respectively. In order to test the robustness of the proposed filter to noise in the reference image, we vary the Std.Dev. of the noise from 5 to 40. For each value, we vary the location of the target in the scene and generate different noise realizations for the both images. Implementation of the $\mathrm{GOF_{AN}}$ requires knowing the target shape and intensity values. Since they are assumed unknown, we estimate them from the reference image. Using the known statistics of the noise in the reference image we use a Wiener filter to attenuate the noise. Thus, we reduce the effects of the noise in the target area and estimate the target shape by rejecting the background with statistics of attenuated noise.

Figure 2 shows the simulation results. The GOF is the optimal filter when the target intensity values and shape are explicitly known. Filter $\mathrm{GOF_{AN\_I}}$ is the filter derived in Sec. 2, assuming that the target information is known. Filter $\mathrm{GOF_{AN\_E}}$ is the filter using estimations of missing parameters. We test two cases: when the noise in the input scene is low (has a Std.Dev. of 5), shown in Fig. 2a; and when the noise in the input scene is moderate (has a Std.Dev. of 15), shown in Fig. 2b. The GOF performs ideally with a DC value near unity, as expected, since there is no noise and no unknown parameters. It is shown as a bound on the possible performance. The performance of the proposed filters is below that of the GOF, even if all information is available. That is because both the $\mathrm{GOF_{AN}}$ variants have a noisy component $R^*(\omega)$. The performance of the $\mathrm{GOF_{AN\_E}}$ filter is high enough to carry out an accurate detection.

(a) Target (b) Background

Figure 1: Images used in the experiments



(a) Performance with low scene noise (b) Performance with moderate scene noise

Figure 2: Computer simulation results

4 Conclusions

In this paper, we proposed an optimum correlation filter for a signal model that considers the presence of additive noise in the reference image and a disjoint cluttered background in the input scene. Parameters required for the filter implementation can be estimated from the observed images. It was shown that the proposed filter along with the proposed estimations is able to detect a noisy target even when the target intensity values and its shape are not explicitly known. The performance of the filter is not significantly affected by a particular realization of the reference image noise.

References

[1] Aguilar-Gonzalez, P.M. and Kober, V. (2008). Pattern recognition of an implicitly given target. *Proceedings of SPIE, Applications of Digital Image Processing XXXI.* pp. 707303.

[2] Javidi, B. (1994). *Real-Time Optical Information Processing.* Academic Press.

[3] Kumar, B.V., Mahalanobis, A., and Juday, R. (2005). *Correlation pattern recognition.* Cambridge Univ Press.