LECTURE 4.2

Target location in clutter

4.2.1 Problem formulation, localization quality criterion and localization device

In this lecture we discuss the problem of target location in images that contain, besides the target object, a clutter of false objects that obcure the target object. Typical examples of such images are images of natural scenes, air and space photographs of the Earth, images in medical diagnostics and computer vision. As it follows from the discussion in Lect. 4.1, it is background clutter objects rather than sensor's noise that represents the main obstacle for reliable object localization in this case. Our purpose therefore is to find out how can one design a localization device that minimizes the danger of false identification of the target object with false objects of the image background.

A general solution of this problem similar to that described in Lect. 4.1 for the model of a single object observed in additive signal independent Gaussian noise does not seem to be feasible in this much more general case for two reasons. First, the least bit constructive statistical or whatever description of image ensemble is very problematic due to great variability of real life images. Second, and this is even more important, in applications where target location in images with cluttered background is required one is interested in the most reliable localization in each particular individual image rather then on average over a hypothetical image ensemble. Consequently, localization algorithm optimization and adaptation for individual images are desired. Obviously, if such an optimization is possible it will provide the best solution for any image ensemble as well.

This requirement of adaptivity will be imperative in our discussion. The second imperative requirement will be that of low computational complexity of localization algorithms. Bearing this in mind we shall restrict the discussion by the same type of localization devices as those described in Lect. 4.1 that consist of a linear filter followed by a unit for locating signal maximum at the filter output (Fig. 2.1.1). Due to the existence of fast and recursive algorithms of digital linear filtering ([9]), such devices have low computational complexity in computer implementation.

They also have very elegant optical implementations in optical correlators ([1,3]). In such devices, it is the linear filter that has to be optimized and is the subject of adaptation.

We will begin the discussion with a formulation of the notion of optimality. Let us assume that the image under analysis can be decomposed into S fragments of area S_s , $s = 1,\ldots,S$ that can be regarded homogeneous in terms of the price for false localization errors (Fig.(2.1.2). Let also $h_s(b/(x_0,y_0))$ be a histogram of image signal amplitudes $b(x,y)$ at the filter output as measured for the s-th fragment over points not contained within the object (background points). We shall assume also that the histogram is measured for the fixed background, and fixed sensor's (imaging system) noise realization and that the object's coordinates are (x_0,y_0), and b_0 is the filter output in the target object location (it may be assumed that $b_0 > 0$ without restricting generality).

The localization device under consideration bases its decision regarding the coordinates of the desired object upon the position of the absolute maximum at the signal output. Consequently, the integral

$$P_{a,s} = \mathbf{AV}_{\mathbf{bg}} \int_{b_0}^{\infty} h_s\left(b/(x_0,y_0)\right)db\,, \qquad (4.\ 2.1.1)$$

with $\mathbf{AV}_{\mathbf{bg}}$ denoting the average over the possible realizations of the background component of the image, represents the portion of the s-th fragment points that, on average, can be erroneously assigned by the decision unit to the target object.

Generally speaking, b_0 should be regarded a random variable as well because it also depends on video signal sensor noise, photographic environment, illumination, object orientation, neighboring objects, and other stochastic factors. In order to take them into consideration, introduce a priori probability density $p(b_0)$ of b_0. Unknown object coordinates also should be regarded random. Moreover, the weight of the measurement errors may differ over different images fragments. To allow for these factors, we introduce weighting functions $w_s(x_0,y_0)$ and W_s characterizing the a priori significance of localization within the s-th fragment and for each s-th fragment, respectively :

$$\iint_{S_s} w_s(x_0, y_0)\,dx_0\,dy_0 = 1\,; \qquad (4.\ 2.1.2)$$

$$\sum_{s=1}^{S} W_s = 1. \qquad (4.2.1.3)$$

Then the quality of estimating coordinates by the device of Fig. 2.1.1 may be described by a weighted mean with respect to $p(b_0)$, $\{w_s(x_0, y_0)\}$, and $\{W_s\}$ of the integral of Eq.(4.2.1.1):

$$\overline{P}_a = \int_{-\infty}^{\infty} p(b_0)\,db_0 \sum_{s=1}^{S} W_s \iint_{S_s} w_s(x_0, y_0) \left(\mathrm{AV}_{bg} \int_{b_0}^{\infty} h_s\big(b/(x_0, y_0)\big)\,db \right) dx_0\,dy_0 \qquad (4.2.1.4)$$

A device that provides the minimum to $\overline{P}_a$ will be regarded optimal on average with respect to background components of the image. If one needs a device optimal for the fixed background part of the image, one should eliminate averaging $\mathbf{AV_{bg}}$ over the background from the above formula.

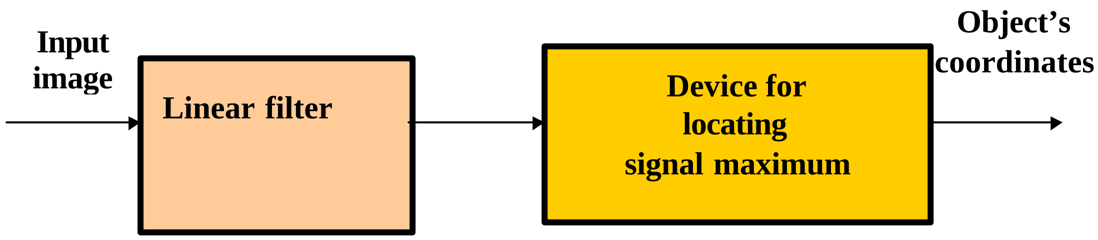


Fig. 4.2.1 Block diagram of the localization device

Fig. 4.2.1 Block diagram of the localization device

4.2.2 Localization of a Precisely Known Object: Spatially Homogeneous Optimality Criterion

4.2.2.1 Ideal optimal adaptive filter

Assume that the target object is defined precisely. In this context it means that the response of any filter to this object may be determined exactly and, therefore, probability density $p(b_0)$ is the delta-function :

$$p(b_0) = \delta\left(b_0 - \bar{b}_0\right). \qquad (4.2.2.1)$$

The Eq. (4.2.1.4), which defines the localization quality, then becomes:

$$\bar{P}_a = \sum_{s=1}^{S} W_s \iint_{S_s} w_s(x_0, y_0)\left(\mathrm{AV}_{bg} \int_{\bar{b}_0}^{\infty} h_s\left(b/(x_0, y_0)\right)db\right)dx_0 dy_0 \, . \qquad (4.2.2.2)$$

If the histogram averaged within each fragment over (x_0, y_0) is denoted by

$$\bar{h}_s(b) = \iint_{S_s} w_s(x_0, y_0) h_s\left(b/(x_0, y_0)\right)dx_0 dy_0 \, , \qquad (4.2.2.3)$$

we obtain:

$$\bar{P}_a = \sum_{s=1}^{S} W_s \left(\mathrm{AV}_{bg} \int_{\bar{b}_0}^{\infty} \bar{h}_s(b)db\right) \qquad (4.2.2.4)$$

Suppose that the optimality criterion is spatially homogeneous, that is, the weights $\{W_s\}$ do not depend on s and are equal to $1/S$. Then

$$\bar{h}(b) = \frac{1}{S}\sum_{s=1}^{S} \bar{h}_s(b) \qquad (4.2.2.5)$$

is the histogram of the filter output signal amplitudes as measured over the whole image and averaged with respect to the unknown coordinates of the target object. By substituting Eq.(4.2.2.5) into Eq.(4.2.2.4), we obtain

$$\bar{P}_a = \int_{\bar{b}_0}^{\infty} \mathrm{AV}_{bg}\left(\bar{h}(b)\right)db \qquad (4.2.2.6)$$

Let us now determine the frequency response $H(f_x, f_y)$ of a filter that minimizes $\overline{P}_a$. The choice of $H(f_x, f_y)$ affects both $\overline{b}_0$ and histogram $\overline{h}(b)$. As $\overline{b}_0$ is the filter response at the object position, it may be determined through the object Fourier spectrum $a(f_x, f_y)$ as

$$\overline{b}_0 = \int_{-\infty}^{\infty} a(f_x, f_y) H(f_x, f_y) df_x df_y \qquad (4.2.2.7)$$

As for the relationship between $\overline{h}(b)$ and $H(f_x, f_y)$, it is, generally speaking, of an involved nature. The explicit dependence of $\overline{h}(b)$ on $H(f_x, f_y)$ may only be written for the second moment m_2^2 of the histogram $\overline{h}(b)$ using Parseval's relation for the Fourier transform:

$$m_2^2 = \iint_S w(x_0, y_0) dx_0 dy_0 \int_{-\infty}^{\infty} b^2 h(b/(x_0, y_0)) db$$

$$= \frac{1}{S_{bg}} \iint_S w(x_0, y_0) dx_0 dy_0 \int_{S_{bg}} b^2(x, y) dx dy$$

$$= \frac{1}{S_{bg}} \iint_S w(x_0, y_0) dx_0 dy_0 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left| a_{bg}^{0,0}(f_x, f_y) \right|^2 \left| H(f_x, f_y) \right|^2 df_x df_y$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \overline{\left| a_{bg}^{0,0}(f_x, f_y) \right|^2} \left| H(f_x, f_y) \right|^2 df_x df_y \,, \qquad (4.2.2.8)$$

where S_{bg} is the area of the background portion of the images, or the image search area minus the area occupied by the signal of the target object under search, $a_{bg}^{0,0}(f_x, f_y)$ is Fourier spectrum of this background portion, and

$$\overline{\left| a_{bg}^{0,0}(f_x, f_y) \right|^2} = \frac{1}{S_{bg}} \iint_S w(x_0, y_0) \left| a_{bg}^{0,0}(f_x, f_y) \right|^2 dx_0 dy_0 \,. \qquad (4.2.2.9)$$

Therefore, we can proceed only by relying upon the Tchebyshev's inequality that connects the probability that a random variable x exceeds some threshold b_0 and the variable's mean value $\overline{x}$ and standard deviation s:

$$\textbf{Probability}\left(x \geq \bar{x} + b_0\right) \leq \sigma^2 / b_0^2 . \qquad (4.2.2.10)$$

Applying this relationship to Eq.(4.2.2.6), we can write:

$$\bar{P}_a = \int_{b_0}^{\infty} \mathrm{AV}_{bg}\left(\bar{h}(b)\right) db \leq \mathrm{AV}_{bg} \frac{m_2^2 - \bar{b}^2}{\left(\bar{b}_0 - \bar{b}\right)^2} , \qquad (4.2.2.11)$$

where $\bar{b}$ is mean value of the histogram $\bar{h}(b)$. By virtue of the properties of the Fourier Transform the latter can be found as:

$$\bar{b} = \frac{1}{S_{bg}} \iint_S w(x_0, y_0) a_{bg}^{0,0}(0,0) H(0,0) dx_0 dy_0 . \qquad (4.2.2.12)$$

It follows from this equation that the mean value of the histogram over the background part of the picture, $\bar{b}$, is determined by the filter frequency response $H(0,0)$ at the point $\left(f_x = 0, f_y = 0\right)$. The same value affects the mean value of the signal over the entire picture being filtered; that is, it defines a constant bias of the signal at the filter output which is irrelevant for the device that localizes the signal maximum in our estimator. Therefore, one can choose $H(0,0) = 0$ and, hence, disregard $\bar{b}$ in (4.2.2.11). Then we can conclude that in order to make the rate $\bar{P}_a$ of anomalous errors minimal one should design the filter as follows:

$$H(f_x, f_y) = \arg\min_{H(f_x, f_y)} \left\{ \mathrm{AV}_{ims} \mathrm{AV}_{bg} \left(\frac{m_2^2}{\bar{b}_0^2} \right) \right\}$$

$$= \arg\min_{H(f_x, f_y)} \left\{ \mathrm{AV}_{bg} \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \overline{\left|a_{bg}^{0,0}(f_x, f_y)\right|^2} \left|H(f_x, f_y)\right|^2 df_x df_y}{\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} a(f_x, f_y) H(f_x, f_y) df_x df_y \right)^2} \right\} \qquad (4.2.2.13)$$

The solution of this equation :

$$H_{opt}\left(f_x, f_y\right) = \frac{\alpha^*\left(f_x, f_y\right)}{\overline{\mathrm{AV}_{\mathrm{bg}}\left|\alpha_{bg}^{0,0}\left(f_x, f_y\right)\right|^2}}, \qquad (4.2.2.14)$$

where asterisk * denotes complex conjugate, follows from Schwarz's inequality :

$$\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \alpha\left(f_x, f_y\right) H\left(f_x, f_y\right) df_x df_y\right)^2$$

$$= \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\alpha\left(f_x, f_y\right)}{\left(\overline{\mathrm{AV}_{\mathrm{bg}}\left|\alpha_{bg}^{0,0}\left(f_x, f_y\right)\right|^2}\right)^{1/2}} H\left(f_x, f_y\right)\left(\overline{\mathrm{AV}_{\mathrm{bg}}\left|\alpha_{bg}^{0,0}\left(f_x, f_y\right)\right|^2}\right)^{1/2} df_x df_y\right)^2$$

$$\le \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\left|\alpha\left(f_x, f_y\right)\right|^2}{\overline{\mathrm{AV}_{\mathrm{bg}}\left|\alpha_{bg}^{0,0}\left(f_x, f_y\right)\right|^2}} \left|H\left(f_x, f_y\right)\right|^2 \left(\overline{\mathrm{AV}_{\mathrm{bg}}\left|\alpha_{bg}^{0,0}\left(f_x, f_y\right)\right|^2}\right) df_x df_y . \qquad (4.2.2.15)$$

One can see that filter (4.2.2.14) is analogous to the optimal filter (1.4.2) described in Lect. 4.1 for object location in correlated Gaussian noise. Power spectrum $\overline{\mathbf{AV}_{\mathbf{bg}}\left|\alpha_{bg}^{0,0}\left(f_x, f_y\right)\right|^2}$ averaged over an ensemble of the background image components plays here the role of additive Gaussian noise power spectrum. There is, however, an important difference. The design of the optimal filter (4.2.2.14) assumes knowledge of averaged power spectrum of the background component of the image $\overline{\mathbf{AV}_{\mathbf{bg}}\left|\alpha_{bg}^{0,0}\left(f_x, f_y\right)\right|^2}$. With the averaging, the filter will be optimal on average over different realizations background image components of the image with the same target object. In applications, however, one usually needs a filter optimal for the particular image rather then optimal on average over a set of images. This means that the averaging over the background image components $\mathbf{AV}_{\mathbf{bg}}$ must be removed from the optimality criterion (4.2.2.13) as it was mentioned in Sect. 2.1. In this way we arrive at the following filter:

$$H_{opt}\left(f_x, f_y\right) = \frac{\alpha^*\left(f_x, f_y\right)}{\left|\alpha_{bg}^{0,0}\left(f_x, f_y\right)\right|^2} \qquad (4.2.2.16)$$

This filter if it can be implemented is optimal for the particular observed image. Its frequency response depends on the image background component and, therefore, it is adaptive. We will call this filter "***ideal optimal adaptive filter***". For the design of this filter one needs to determine the power spectrum of the background component of the image from the observed image that contains also the target object. Because coordinates of the target object are not known one cannot separate the target object and the background in the observed image and therefore can not exactly determine the background component power spectrum needed for the filter design. The ideal optimal adaptive filter can not be implemented in practice. One can however attempt to approximate it by means of an appropriate estimation of the background component power spectrum from the observed image. Such approximate filters will be called ***optimal adaptive filters***.

4.2.2.2 Estimation of the background image component power spectrum for the optimal adaptive filter design

In this section, we will discuss methods of estimating the background component power spectrum from the observed image for the design of optimal adaptive filters. One can suggest two models of how the target object is represented in the image: an additive model and an implant model.

In the additive model, the observed image $\boldsymbol{b}(x,y)$ is regarded as an additive composition of the target signal $\boldsymbol{a}(x-x_0, y-y_0)$ in (unknown) coordinates (x_0, y_0) and the background component $\boldsymbol{a}_{bg}(x,y)$:

$$\boldsymbol{b}(x,y) = \boldsymbol{a}(x-x_0, y-y_0) + \boldsymbol{a}_{bg}(x,y) \qquad (4.2.2.17)$$

One can show that ([6]), in the assumption of uniform distribution of the target coordinates over the image area, the background component power spectrum can be approximated as

$$\overline{\left|\boldsymbol{a}_{bg}(f_x, f_y)\right|^2} \cong \left|\boldsymbol{b}(f_x, f_y)\right|^2 + \left|\boldsymbol{a}(f_x, f_y)\right|^2 \qquad (4.2.2.18)$$

which implies the following optimal adaptive filter:

$$\boldsymbol{H}^{a}_{opt}(f_x, f_y) = \frac{\boldsymbol{a}^*(f_x, f_y)}{\left|\boldsymbol{b}(f_x, f_y)\right|^2 + \left|\boldsymbol{a}(f_x, f_y)\right|^2} \qquad (4.2.2.19)$$

Implant model assumes that

$$a_{bg}(x,y) = [1 - w(x - x_0, y - y_0)]b(x,y), \qquad (4.2.2.20)$$

where $w(x - x_0, y - y_0)$ is a target object window function:

$$w(x - x_0, y - y_0) = \begin{cases} 1, & \text{within the target object} \\ 0, & \text{elsewhere} \end{cases}, \qquad (4.2.2.21)$$

For this model, the following approximation of the background component power spectrum can be used in the assumption of uniform distribution of the target co-ordinates over the image area ([6]):

$$\overline{\left|a_{bg}(f_x, f_y)\right|^2} \cong \left(1 - \frac{2S_{ob}}{S_{img}}\right)\left|b(f_x, f_y)\right|^2 + \frac{S_{obj}}{S_{img}}\left|b(p_x, p_y)\right|^2 \cdot \left|W(f_x, f_y)\right|^2 \qquad (4.2.2.22)$$

where S_{obj} and S_{img} are area occupied by the image and target object, respectively, $W(f_x, f_y)$ is Fourier transform of the object window function $w(x,y)$ and $\cdot$ denotes convolution. Therefore the adaptive optimal filter for implant model takes form:

$$H_{opt}^{impl}(f_x, f_y) \cong \frac{a^*(f_x, f_y)}{\left(1 - \frac{2S_{obj}}{S_{img}}\right)\left\{\left|b(p_x, p_y)\right|^2\right\} + \frac{S_{ob}}{S_{img}}\left[\left|b(p_x, p_y)\right|^2\right] \cdot \left|W(f_x, f_y)\right|^2}. \qquad (4.2.2.23)$$

When, as it is usually the case, the area occupied by the target object is much smaller than that of the image ($S_{ob} << S_{img}$), a zero order approximation

$$\overline{\left|a_{bg}(f_x, f_y)\right|^2} \cong \left|b(f_x, f_y)\right|^2 \qquad (4.2.2.24)$$

to the background image component power spectrum can be used for both model which results in the following filter:

$$H_{opt}^{0}(f_x, f_y) = \frac{a^*(f_x, f_y)}{\left|b(p_x, p_y)\right|^2} \qquad (4.2.2.25)$$

Experimental experience shows also ([6]) that the filter obtained for background image spectrum estimation by input image spectrum smoothing provides good results

$$H_{opt}^{smth}(f_x, f_y) \cong \frac{a^*(f_x, f_y)}{\left[\left|b(p_x, p_y)\right|^2\right] \cdot W(f_x, f_y)} \qquad (4.2.2.26)$$

In the conclusion of this section we illustrate advantages of the optimal adaptive correlator with respect to the conventional matched filter correlator that is optimal for target location in Gaussian noise and in absence of false objects and clutter. Figs. 4.2.2.1 and 4.2.2.2 present two typical examples. Fig. 4.2.2.1 demonstrates localization of a fragment of right of two stereoscopic images on the left one. In Fig. 4.2.2.2 results are presented of an experiment in which a round spot of about 32 pixel in diameter (Fig. 4.2.2.2, c) imitating a tumour was randomly placed within a chest X-ray image. For different contrast of the spot, average rate of false localization was then found over an ensemble of the spot random positions. From plots in in Figs. 4.2.2.1 and 4.2.2.2 one can see how much the optimal adaptive filter suppresses signal from background image component thus reducing probability of false detection. In particular, plots in Fig.4.2.2.2 show that, with the optimal adaptive filter correlator, the spot can be reliably detected for much lower contrast than with the matched filter correlator.

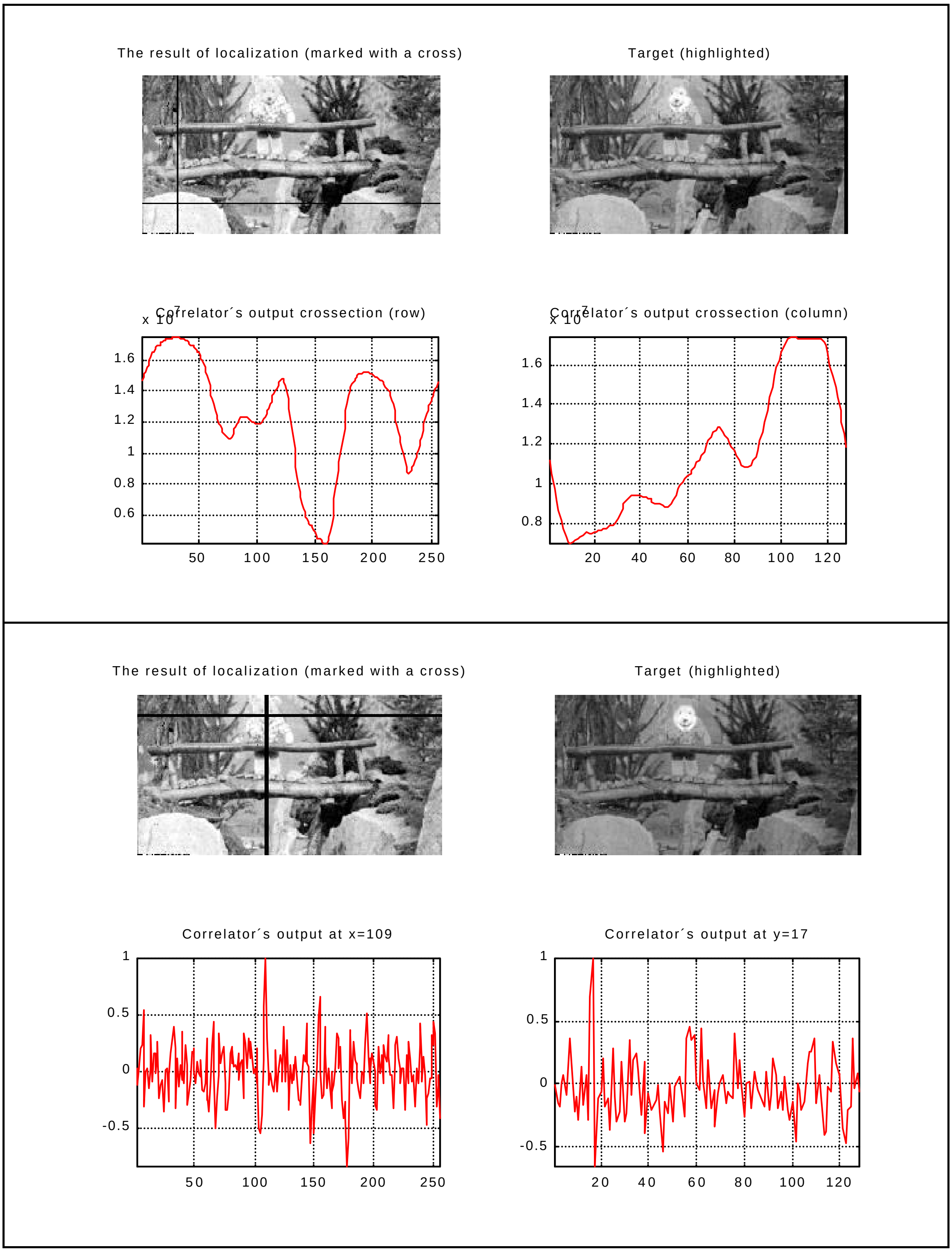
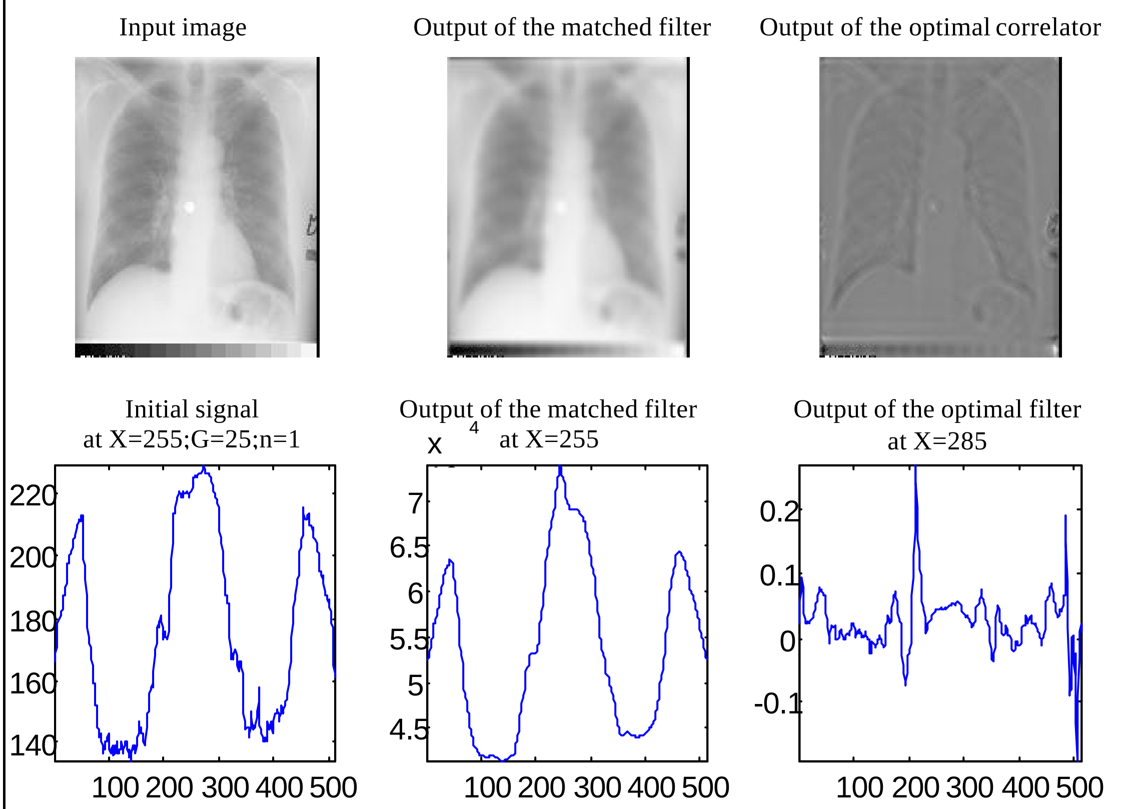
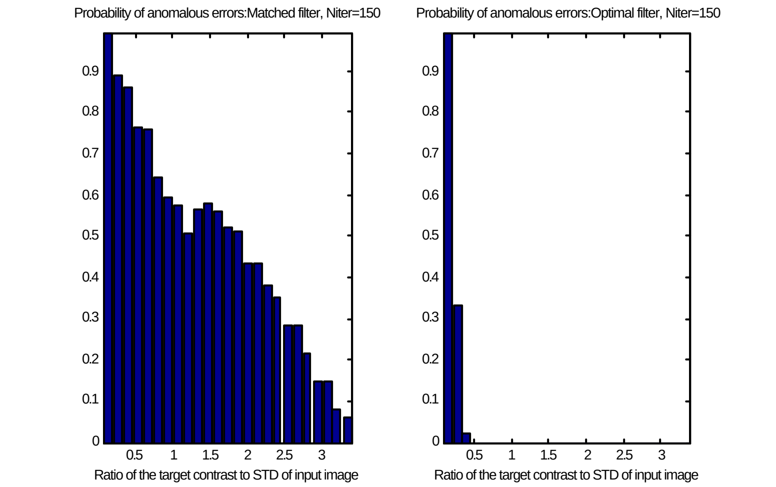


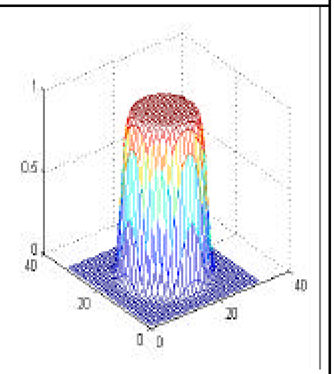
Fig. 4.2.2.1 Comparison of the matched (upper) and optimal adaptive filter (bottom) correlators for localization of correspondent points in stereoscopic images



a) Initial test image and outputs of the matched filter and optimal correlatort



b) Probability of false detection as a function of the spot contrast



c)

3_d plot of the model spot

Fig. 4.2.2.2 Comparison of the matched filter and optimal adaptive correlators capability of locating a tumour modelled as a round spot randomly placed within a chest X-ray image

4.2.3 Localization of Inexactly Known Objects: Spatially homogeneous criterion

In the case of an inexactly defined object there exists uncertainty about object parameters such as orientation, scale, other geometrical distortions, distortions due to sensor's noise, etc. The probability density $p(b_0)$ of the filter response to the target object in its coordinates can no longer be regarded to be the delta function. Therefore the filter of the optimal estimator must provide the minimum of the integral:

$$\overline{P}_a = \int_{-\infty}^{\infty} p(b_0) \int_{b_0}^{\infty} \mathrm{AV}_{bg}\left(\overline{h}(b)\right) db$$

(4.2.3.1)

where $\overline{h}(b)$ is defined by Eq.(4.2.2.3).

In this case, two options should be considered, depending on implementation restrictions.

a. *Localization device with selection*. Decompose the interval of possible values of b into subintervals within which $p(b_0)$ may be regarded as constant. Then

$$\overline{P}_a = \sum_k p_k \int_{\overline{b}_0^{(k)}}^{\infty} \mathrm{AV}_{bg}\left(\overline{h}(b)\right) db \qquad (4.2.3.2)$$

where $b_o^{(k)}$ is a representative of the k-th interval and p_k is the area under $p(b_0)$ over the k-th interval. Because $p_k \geq 0$, $\overline{P}_a$ is minimal, if

$$\overline{P}_a^{(k)} = \int_{\overline{b}_0^{(k)}}^{\infty} \mathrm{AV}_{bg}\left(\overline{h}(b)\right) \qquad (4.2.3.3)$$

is minimal. The problem, thus, reduces to the above problem of localizing a precisely known object, the only difference being that now there are several given objects, and, therefore, the optimal filter

$$H_{opt}^{(k)}(f_x, f_y) = \frac{\alpha^{*(k)}(f_x, f_y)}{\mathrm{AV}_{ims}\mathrm{AV}_{bg}\overline{\left|\alpha_{bg}^{0,0}(f_x, f_y)\right|^2}} \qquad (4.2.3.4)$$

needs to be generated separately for each "representative" of all possible object variations. Of course, generating and implementing multiple filters as well as the task of filtering itself requires additional time and hardware. If this is unacceptable, the only alternative is to adjust the filter to an averaged object.

b . *Localization device adjusted to an averaged object*. If the variance of the object's parameters is not too large, one may, at the expense of a higher anomalous error rate, solve the problem as though the object were exactly known. The optimal filter in this case is to be corrected with due regard to the object parameter dispersion. In order to show this, change the variables $\beta_1 = \beta - \beta_0$ and change the order of integration in Eq.(4.2.3.1)

$$\overline{P}_a = \int_0^{\infty} d\beta_1 \int_{-\infty}^{\infty} p(\beta_0) \mathrm{AV}_{bg}\left(\overline{h}\left(\beta_1 + \overline{\beta}_0\right)\right) d\beta \qquad (4.2.3.5)$$

The internal integral in Eq.(4.2.3.5), is a convolution of distributions, or distribution of the difference of two variables β and β_0. Denote this distribution by $\overline{h}_p(\beta_1)$. Its mean value is equal to the difference of mean values $\overline{\beta}_0$ and β_{av} of the distributions $p(\beta_0)$ and $h(\beta)$, respectively. Further, its variance is equal to the sum of the variances of these distributions; that is, $\left[\mu_2^2 - (\beta_{av})^2\right] + \sigma_\beta^2$, where σ_β^2 is the variance of the distribution $p(\beta_0)$. Therefore,

$$\overline{P}_a = \int_0^{\infty} \overline{h}_p(\beta_1) d\beta_1 = \int_{\overline{\beta}_0}^{\infty} \mathrm{AV}_{bg}\left(\overline{h}_p\left(\beta_1 - \overline{\beta}_0\right)\right) d\beta_1 \qquad (4.2.3.6)$$

Thus, the problem has again been boiled down to that of Lect.4.2.2, and one can use Eq.(4.2.2.14) as the optimal filter frequency response with appropriate modifications of its numerator and denominator. As $\overline{\beta}_0$ is a mean value of the filter output at the object's location over the distribution $p(\beta_0)$, a complex conjugate spectrum $\alpha^*(f_x, f_y)$ of the object in the numerator of the formula (4.2.2.14) should be replaced by a complex conjugate of object spectrum $\overline{\alpha}^*(f_x, f_y)$ averaged over variations of the target object. As for the denominator of Eq.(4.2.2.14), it should be modified by adding the variance of the object spectrum

$$\left|\alpha_{ef}(f_x, f_y)\right|^2 = \mathrm{AV}_{obj}\left|\alpha(f_x, f_y) - \overline{\alpha}(f_x, f_y)\right|^2 . \qquad (4.2.3.7)$$

Therefore, we obtain the following expression for the optimal-in-average filter frequency response :

$$\overline{H}_{opt}(f_x, f_y) = \frac{\overline{\alpha}^*(f_x, f_y)}{\mathrm{AV}_{bg}\overline{\left|\alpha_{bg}^{0,0}(f_x, f_y)\right|^2} + \left|\alpha_{ef}(f_x, f_y)\right|^2} \qquad (4.2.3.8)$$

As before, this filter will be optimal-in-average over all variations of the object and all possible backgrounds. By eliminating averaging over background we obtain an adaptive filter :

$$\overline{H}_{opt}(f_x, f_y) = \frac{\overline{\alpha}^*(f_x, f_y)}{\overline{\left|\alpha_{bg}^{0,0}(f_x, f_y)\right|^2} + \left|\alpha_{ef}(f_x, f_y)\right|^2} \qquad (4.2.3.9)$$

which will be optimal-in-average over variations of the object but for fixed background; that is for the given observed picture. For the design of this filter, methods of estimating background image component power spectrum $\overline{\left|\alpha_{bg}^{0,0}(f_x, f_y)\right|^2}$ discussed in Lect.4.2.2 can be used. In a special case when variations of the target object are caused by additive sensor's noise of the imaging system, the variance of the object spectrum is equal to the spectral density of the noise $\overline{\left|\nu(f_x, f_y)\right|^2}$, and optimal filter becomes:

$$\overline{H}_{opt}(f_x, f_y) = \frac{\overline{\alpha}^*(f_x, f_y)}{\overline{\left|\alpha_{bg}^{0,0}(f_x, f_y)\right|^2} + \overline{\left|\nu(f_x, f_y)\right|^2}} \qquad (4.2.3.10)$$

As it was already mentioned, such an averaged filter cannot provide as large "signal-to-noise" ratio as the filter adjusted for a precisely known object. Indeed, for each specific object with spectrum $\alpha(f_x, f_y)$ the following inequality for "signal-to-noise ratio" $\overline{SNR}$ for the filter (4.2.3.9) holds:

$$\overline{SNR} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\overline{\alpha^*(f_x, f_y)}\,\overline{\alpha}^*(f_x, f_y)}{\overline{\left|\alpha_{bg}^{0,0}(f_x, f_y)\right|^2} + \left|\alpha_{ef}(f_x, f_y)\right|^2} df_x df_y$$

$$\leq \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\left|\alpha(f_x, f_y)\right|^2}{\overline{\left|\alpha_{bg}^{0,0}(f_x, f_y)\right|^2} + \left|\alpha_{ef}(f_x, f_y)\right|^2} df_x df_y \qquad (4.2.3.11)$$

The right hand part of the inequality is the "signal-to-noise ratio" for the optimal filter, exactly matched to the object.

In the conclusion of this section consider one important special case of the target object variability that has to be taken into account when target object can be regarded as implanted in the background. This is the variability due to the background component of the image. If the target object is implanted into the background, the target object signal component in the observed image can be described as

$$\boldsymbol{a}(x,y) = \boldsymbol{a}_0(x,y) - \boldsymbol{wnd}(x,y)\boldsymbol{b}(x - x_0, y - y_0), \qquad (4.2.3.12)$$

where $\boldsymbol{wnd}(x,y)$ is a target object window

$$\boldsymbol{wnd}(x,y) = \begin{cases} 1, & \text{within the target object} \\ 0, & \text{elsewhere} \end{cases}, \qquad (4.2.3.13)$$

$\boldsymbol{a}_0(x,y)$ is a “pure” target object signal assumed zero outside the target object location:

$$\boldsymbol{a}_0(x,y) = \boldsymbol{a}_0(x,y)\boldsymbol{wnd}(x,y) \qquad (4.2.3.14)$$

and $\boldsymbol{b}(x,y)$ is the background image. In this representation, the variations of the observed target object are defined by the term $\boldsymbol{wnd}(x,y)\boldsymbol{b}(x - x_0, y - y_0)$. Therefore,

$$AV_{obj}\left(\boldsymbol{a}(x,y)\right) = \boldsymbol{a}_0(x,y) - AV_{obj}\left(\boldsymbol{wnd}(x,y)\boldsymbol{b}(x - x_0, y - y_0)\right), \qquad (4.2.3.15)$$

or, in the spectral domain,

$$\begin{aligned} &AV_{obj}\left(\boldsymbol{a}(f_x, f_y)\right) = \\ &\boldsymbol{a}_0(f_x, f_y) - \mathrm{AV_{obj}}\left\{\Omega(f_x, f_y) \otimes \left[\boldsymbol{b}_0(f_x, f_y)\exp\left[i2\pi(f_x x_0 + f_y y_0)\right]\right]\right\} = \\ &= \boldsymbol{a}_0(f_x, f_y) - \Omega(f_x, f_y) \otimes \left(\boldsymbol{b}_0(f_x, f_y)CF(f_x, f_y)\right), \qquad (4.2.3.16) \end{aligned}$$

where $CF(f_x, f_y)$ is the characteristic function (Fourier transform) of the distribution density of the object coordinate (x_0, y_0), $\Omega(f_x, f_y)$ is Fourier spectrum of the target window function $\boldsymbol{wnd}(x,y)$ and $\otimes$ designates convolution. If target object coordinates (x_0, y_0) are uniformly distributed over an area of search S_{img} that is much larger than the object size, its distribution characteristic function $CF(f_x, f_y)$ can be approximated by the delta-function, divided by S_{img}, and therefore

$$\mathbf{AV}_{\mathbf{obj}}\left(\boldsymbol{\alpha}\left(f_x, f_y\right)\right) \approx \boldsymbol{\alpha}_0\left(f_x, f_y\right) - \Omega\left(f_x, f_y\right)\cdot \boldsymbol{\beta}_0\left(\mathbf{0}\right)/\, \boldsymbol{\sigma}_{img} \;=,$$

$$\boldsymbol{\alpha}_0\left(f_x, f_y\right) - \Omega\left(f_x, f_y\right)\cdot \overline{\boldsymbol{img}} \qquad (4.2.3.17)$$

where $\overline{\boldsymbol{img}}$ is the arithmetic mean over the background image component.

4.2.4. Localization in the case of a spatially inhomogeneous criterion

Let us return to the general formula, Eq.(4.2.1.14) that accounts, by means of weight coefficients $\{W_s\}$, for possible inhomogeneity of the localization criterion Depending on the implementation constraint, one of three ways to attain the minimum of $\overline{P}_a$ may be suggested in this case.

A. *Readjustable localization device with fragment-wise optimal filtering.* Given nonnegative weights $\{W_s\}$, the minimum of $\overline{P}_a$ is attained at the minima of

$$\overline{P}_s = \int_{b_0} p(b_0)db_0 \iint_{S_s} w_s\left(x_0, y_0\right)\left(\text{AV}_{bg}\int_{b_0}^{\infty} h_s\left(b/\left(x_0, y_0\right)\right)db\right)dx_0dy_0 \qquad (4.2.4.1)$$

This implies that the filter should be re-adjustable for each ***s***-th fragment, filtering being carried out with respect to the fragments within which the averaging in Eq.(4.2.4.1) is done. For each fragment, the optimal filter is determined as it was shown in Sect. 2.3 for homogeneous criterion on the basis of measurements of the observed local spectra of each fragment. Note that the re-adjustable filter response does no depend on the weights $\{W_s\}$ or on the corresponding continuous weight function.

According to Eq.(4.2.1.4), the fragments are assumed not to overlap. It is obvious from the very sense of Eq.(4.2.1.4), that it gives rise to a sliding processing algorithm based upon an estimate of the current local power spectrum of the picture because the error weights $\{W_s\}$ may be defined by a continuous function of image coordinates. Fig. 4.2.4.1 illustrates such a localization with optimal filtering in moving window ([4,5]).

B. *Non-readjustable localization device*. When there is no possibility of implementing the adjustable localization device with fragment-wise or sliding processing , the alternative is to design a localization device adjusted to the power spectra of picture fragments averaged over $\{W_s\}$. Indeed, it follows from Eq.(1.2.4), that

$$\overline{P}_a = \int_{-\infty}^{\infty} p(b_0)\,db_0 \sum_{s=1}^{S} W_s \iint_{S_s} w_s(x_0,y_0)\left[\mathrm{AV}_{\mathrm{bg}} \int_{b_0}^{\infty} h_s\big(b/(x_0,y_0)\big)\,db\right] dx_0\,dy_0$$

$$= \int_{-\infty}^{\infty} p(b_0)\,db_0\, \mathrm{AV}_{\mathrm{bg}} \int_{b_0}^{\infty} \overline{\overline{h}}_s(b)\,db \qquad (4.2.4.2)$$

where $\overline{\overline{h}}_s(b)$ is a histogram averaged over $\{W_s\}$ and $\{w_s(x_0,y_0)\}$:

$$\overline{\overline{h}}_s(b) = \sum_{s=1}^{S} W_s \iint_{S_s} w_s(x_0,y_0)\left[\mathrm{AV}_{\mathrm{bg}} \int_{b_0}^{\infty} h_s\big(b/(x_0,y_0)\big)\,db\right] dx_0\,dy_0 \qquad (4.2.4$$

Whence one may conclude by analogy with Eqs.(4.2.3.9) that

$$\overline{\overline{H}}_{opt}(f_x,f_y) = \frac{\alpha^*(f_x,f_y)}{\mathrm{AV}_S \overline{\left|\alpha_{bg}^{0,0}(f_x,f_y)\right|^2} + \left|\alpha_{ef}(f_x,f_y)\right|^2} \qquad (4.2.4.4)$$

where AV_S means averaging image fragment power spectra over the image area according to weights $\{W_s\}$. Thus, the optimal filter frequency response in this case depends on the weights $\{W_s\}$ and is determined by the averaged power spectrum of all fragments. Of course, such an "averaged" filter may not be as efficient for certain image fragments as the filter adapted to this fragment but it definitely is less expensive computation wise. One can notice also that if a set of images is given for which the filter should be optimized, the averaging should be extended to this set. In the limit, it reduces to estimation of the image statistical power spectrum, and we arrive at the optimal filter (4.1.15) for additive non-white Gaussian noise.

Very frequently one can apply above reasoning to pattern recognition tasks. For instance, in face or finger print recognition one recognize an image from a given set of images. It follows from Eq. (4.2.4.4) that, prior to recognition, images in the set should be normalized and the normalization can be implemented in the Fourier domain by the filter:

$$H_{norm}(f_x,f_y) = \frac{1}{\sum_{k=1}^{N} \left|\alpha_k(f_x,f_y)\right|^2}, \qquad (4.2.4.5)$$

where k is image index and N is number of images in the set. The recognition can then be carried out by matched filtering the normalized images with filters

$$H_k\left(f_x, f_y\right) = \frac{\alpha_k^*\left(f_x, f_y\right)}{\sum_{k=1}^{N}\left|\alpha_k\left(f_x, f_y\right)\right|^2} \qquad (4.2.4.6)$$

and subsequent determining the filter that provides the highest output.

C. *Image homogenization.* In cases when non re-adjustable localization device does not provide admissible localization reliability while the computational expenses for the implementation of the adjustable device one also can not afford one can accept an intermediate solution, an image preprocessing that improves image homogeneity. The idea of image homogenization by preprocessing follows from the following reasoning. In terms of the design of the optimal adaptive filter for the localization device, image can be regarded homogeneous if spectra of all image fragments involved in the averaging procedure in Eq.(4.2.4.4) are identical. While this is, in general, not the case, one can attempt to reduce the diversity of image local spectra by means of image preprocessing. One of the computation-wise simplest preprocessing is the preprocessing that standardizes image local mean and local variance. In discrete signal representation, it is defined for image fragments of $(2N_1+1)(2N_2+1)$ pixels as

$$\tilde{b}(k,l) = \sqrt{\frac{\sigma_{st}^2}{\overline{b^2(k,l)}}}\left[b(k,l) - \bar{b}(k,l)\right] \qquad (4.2.4.7)$$

where (k,l) are running coordinates of the window central pixel,

$$\bar{b}(k,l) = \frac{1}{(2N_1+1)(2N_2+1)}\sum_{n=-N_1}^{N_1}\sum_{m=-N_2}^{N_2} b(k-n, l-m) \qquad (4.2.4.8)$$

is image local mean over the window and

$$\overline{b^2(k,l)} = \frac{1}{(2N_1+1)(2N_2+1)}\sum_{n=-N_1}^{N_1}\sum_{m=-N_2}^{N_2} b^2(k-n, l-m) - \left[\bar{b}(k,l)\right]^2 \qquad (4.2.4.9)$$

is its local variance. Computation of image local mean and variance can be very efficiently carried out recursively such that the computational complexity does not depend on the window size ([3]).

Such an image homogenization procedure equalizes average of local power spectra over all frequencies and can be regarded as a zero order approximation to making them identical on each particular frequency. Selection of the window size is governed by the image inhomogeneity. An example of localization with homogenization preprocessing is shown in Fig. 4.2.4.2.

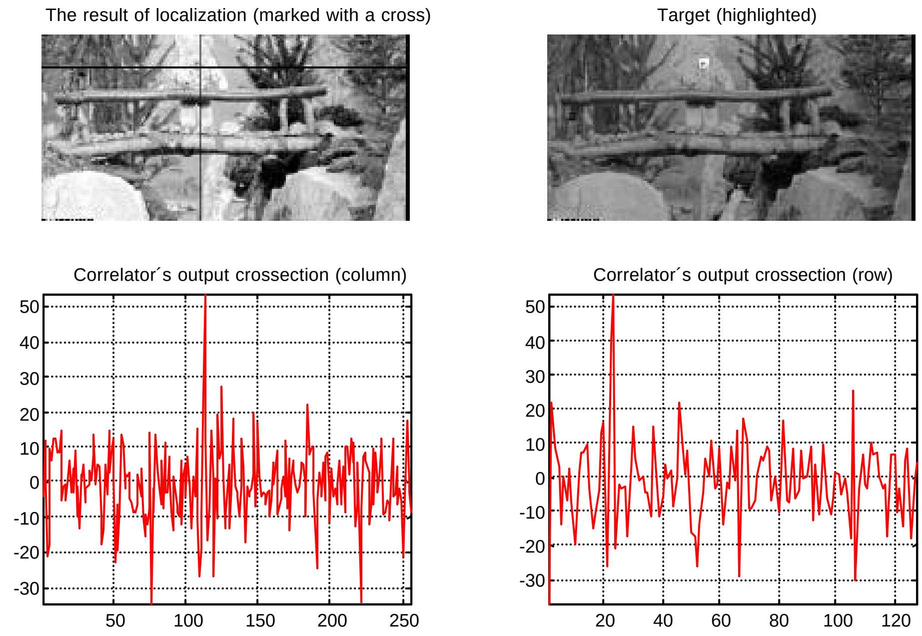


Fig. 4.2.4.1 Localization of a small (8x8 pixels, highlighted) fragment of the right of two stereoscopic images on the left image by the local adaptive filter. Graphs illustrate difference between filter response to the target and that to the rest of the image. Running window size is 32x32 pixels; image size is 128x256 pixels. The same filter applied globally rather than locally fails to locate the fragment.

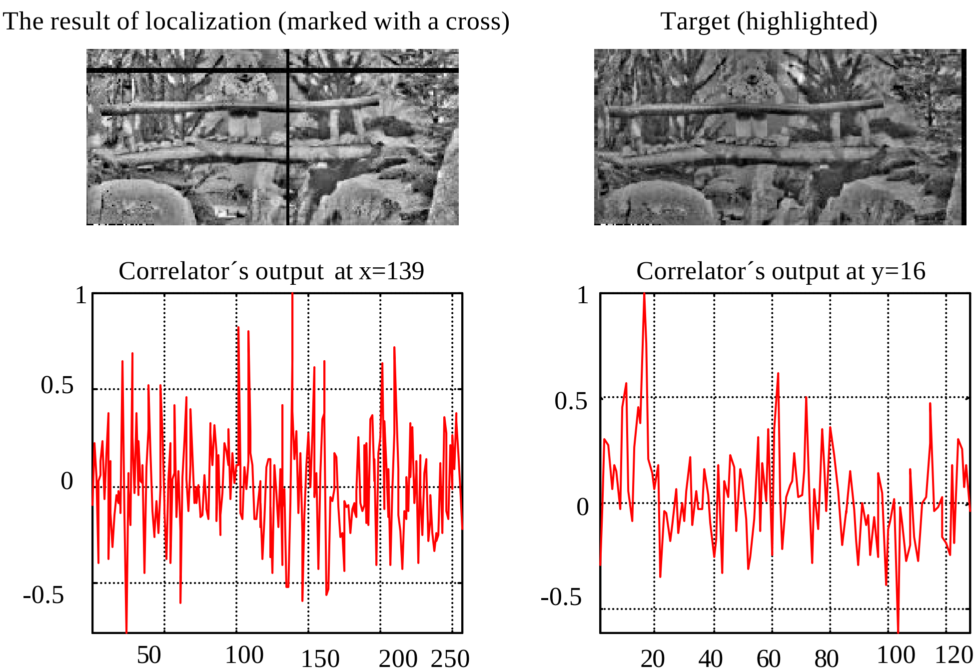


Fig. 4.2.4.2 Illustration of improvement of localization reliability by means of image “homogenization”

References

1. A. Van der Lugt, Signal Detection by Complex Spatial Filtering, IEEE Trans., IT-10, 1964, No. 2, p. 139
2. L.P. Yaroslavsky, The Theory of Optimal Methods for Localization of Objects in Pictures, In: Progress in Optics, Ed. E. Wolf, v. XXXII, Elsevier Science Publishers, Amsterdam, 1993
3. L. Yaroslavsky, M. Eden, Fundamentals of Digital Optics, Birkhauser, Boston, 1996
4. L. Yaroslavsky, Local Adaptive Filtering in Transform Domain for Image Restoration, Enhancement and Target Location, in: 6th Int. Workshop on Digital Image Processing and Computer Graphics (DIP-97), Em. Wenger and L. Dimitrov, eds., 20-22 Oct., 1997, Vienna, Austria, SPIE vol. 3346
5. L. Yaroslavsky, Image Restoration, enhancement and target location with local adaptive filters, in:International Trends in Optics and Photonics, ICOIV, ed. by T.Asakura, Springer Verlag, 1999, pp. 111-127
6. L. Yaroslavsky, Target Location: Accuracy, Reliability and Optimal Adaptive Filters, TICSP series, Tampere Int. Center for Signal processing, TTKK, Monistamo, 1999