This chapter is devoted to one of the characteristic problems of automatic picture interpretation, namely the localization of objects in a picture. A mathematical formulation of the problem is given in Sect. 9.1, and in Sect. 9.2 the optimal adaptive algorithm for localization of a precisely known object in a spatially homogeneous picture is derived and illustrated. Section 9.3 gives modifications of the algorithm for use when there is some uncertainty about the object and when the picture is inhomogeneous and blurred. In Sect. 9.4 the connection between optimal localization and picture contours is discussed, and an algorithm for automatic choice of the most easily detectable objects is suggested and illustrated. In Sect. 9.5, the automatic detection and extraction of benchmarks on aerial and space photographs is presented as a practical application of the methods developed.

9.1 Problem Formulation

One of the most important purposes of pictures is to carry information about the spatial distribution of objects in relation to one another. There are a number of practical tasks which require that objects be found and their coordinates measured (localization). These include detecting and measuring the coordinates of objects in photographic interpretation; measuring the mutual arrangement of corresponding points on stereo pairs of pictures, finding and measuring the position of benchmarks in aerial and space photographs; detecting faults and alien formations in medical and industrial diagnostics using pictures; finding set objects and symbols in information and retrieval systems; etc. These tasks can be fulfilled either "manually" by visual picture analysis, or with automatic digital or analogue (optical or electronic) computers by processing the corresponding 2-D signal.

To solve problems of automatic detection and/or localization (measurement of the coordinates) of objects in pictures, a description of the signal corresponding to the object sought must be available; otherwise the whole task...
becomes meaningless! Moreover, the means whereby this signal is represented in the observed picture must be also described.

The simplest additive model of the object sought and the picture observed is commonly used [9.1,2]. In it the picture observed is considered to be the sum of the signal sought, which is known with an accuracy up to a coordinate shift, and additive, signal-independent normal noise with a known correlation function. This model gives the well-known result that the optimal estimator of the coordinates of the signal sought, which yields the maximum a posteriori probability estimate, will be a linear estimator. The latter is composed of the linear optimal filter and the decision-making device defining the coordinates of the absolute maximum or a number of principal maxima (for the localization of several objects) of the signal at the filter output [9.3,4]. The linear optimal filter has the frequency response

\[ H(f_1, f_2) = \frac{a_0^*(f_1, f_2)}{P_n(f_1, f_2)} \]  

where \( f_1, f_2 \) are the spatial frequencies; \( a_0^*(f_1, f_2) \) is the complex-conjugate of the spectrum of the signal to be found taken at the origin of coordinates; and \( P_n(f_1, f_2) \) is the power spectral density of the additive noise.

If the additive Gaussian noise is "white", i.e., if it has a uniform power spectrum \( P_n(f_1, f_2) = \text{const} \), the optimal filter becomes the matched filter, or correlator [9.4,5]:

\[ H(f_1, f_2) = a_0^*(f_1, f_2) \]  

However, for many practical tasks of localizing objects in pictures the additive model is false, and in general the relationship between the signal sought and the picture against whose background it is to be found cannot be described deterministically. Thus, for example, in aerial photographs, signals from separate objects are not summed in the observed general signal, but "mortice" into it. Moreover, the observed signal of the object sought is determined not only by the object itself, but also by the neighbouring objects (e.g., their shadows), the lighting conditions during filming, weather conditions, camera noise, and other random factors that are difficult to formalize. Therefore the problem of localization objects in pictures should be solved with more realistic assumptions about the form in which the object is represented in the observed picture, and also with the least possible number of limitations.

There is yet another approach to detecting and localizing objects in pictures. It can be termed heuristic, and is characteristic of research on pat-