Special Guest Editorial

Optical Pattern Recognition: An Overview

A. VanderLugt
Harris Government Systems Sector
P.O. Box 37
Melbourne, Florida 32901

During a recent trip I waited for my suitcase after a heavily loaded flight from Atlanta to Los Angeles. As the various baggage items appeared on the conveyor, I reflected on the general problem of pattern recognition. The bags came in various shapes, colors, sizes, and orientations: my eyes casually scanned them as they appeared, and I found that I could quickly dismiss those that did not match my mental image of my own bag. I immediately recognized my bag when it appeared (it does have some distinctive scars after years of service) and resumed my trip to the hotel. While I was waiting, I also found time to observe the people gathered in the baggage area to see if there was anyone I recognized. Again, I was reminded of how subtle differences in facial expression, posture, and general body style allow us to instantly recognize people we know.

The eye-brain combination is a marvelous processor for pattern recognition, and much effort has been dedicated to finding automated or semiautomated means for approaching, even in a small way, its performance. Other species, with less highly developed eye-brain skills, are also able to recognize objects vital to their survival; however, it is not so apparent how this is done. Species with segmented lenses, for example, each with rather poor resolution, are able to recognize objects; a more thorough study of this process may help us to understand more fully the necessary conditions for pattern recognition.

The recognition of three-dimensional objects, independent of scale, orientation, and spatial position, is very important in the emerging field of robotics. Early investigators of pattern recognition focused on the simpler problem of character recognition; here, the scale, orientation, and font can be controlled, leaving the factor of spatial position as the main parameter to be estimated. An algorithm in which the unknown character is compared, or correlated, with those stored on a spatial mask is called template matching. A high degree of correlation, of course, denotes a high degree of similarity to one of the stored symbols. The total process, however, took considerable time because all spatial locations of the test character could not be searched in parallel.

Coherent optical processing offers a solution to parallel searching for all possible positions, provided that the template can be realized as a matched filter in the Fourier domain. Some of the impetus for using matched filtering as a first-order approximation to the human recognition process arose from its successful use in radar systems. The possibility of implementing the filter in the Fourier domain is a unique attribute of optical systems; this fact, coupled with the inherent two-dimensional image processing capability, spurred efforts to discover ways to construct an optical matched filter.

The matched filter for an arbitrary object shape contains both amplitude and phase information. The amplitude is easily recorded on photographic film, but recording the phase is a more difficult problem. If the phase response is fairly simple, such as that required for a signal whose spectrum is real, it can be generated independently of the amplitude response through the use of deposition or etching of thin films. In 1964 I reported a method, closely related to that of holography, for constructing a matched filter in which the necessary amplitude and phase responses can be recorded on photographic film without explicitly computing the Fourier transform of the signal. It then became possible to perform pattern recognition with the entire object field being searched in parallel and with surprisingly good results.

A somewhat more difficult problem is that of recognizing objects of military interest in imagery. Here, a whole set of uncontrollable parameters affects the results. Imagery represents the two-dimensional projection of three-dimensional objects. As a result, the recorded signals change as a function of the orientation of the object as well as of the viewing angle. The scale of the signals is often unknown or may change within a scene. The recorded signal is influenced by the sun angle in photography, the head distribution in infrared imagery, or the specular reflectivity in radar. We still are faced with recognizing *objects*, but the *patterns* that represent the objects now have significantly more unknown parameters. Also, the signal-to-noise ratio is generally much less in imagery in comparison to, say, that of a document containing characters.

Nevertheless, some impressive results were obtained by early investigators using matched filtering on reconnaissance imagery. Although the background noise statistics are far from the white, Gaussian noise on which the matched filtering model is based, prewhitening filters based on average scene statistics were used to implement the more generalized optimum matched filter. Filters having a degree of tolerance to scale, orientation, and other parameter uncertainties were constructed by controlling the frequency response.

Much effort has been expended since the early and middle 1960s to improve the performance and utility of optical pattern recognition, and I will now trace some of the threads of these efforts.

The matched filter, as well as other filtering operations, also can be constructed by calculating the complex-valued Fourier transform digitally and using a plotter to record the desired function. These computer-generated filters present an opportunity to more finely control the response of the filter and also pave the way for the capability to electronically change the filter in situ. Generally, the time-bandwidth product is somewhat reduced when using these schemes because the desired analog response is obtained through the use of binary encoding.

Another thread of activity has been the development of spatial light modulators to replace the photographic film used both for recording the input imagery and the matched filter. Such devices, in which the transmittance can be controlled electrically, are desirable because many modern sensors are based on CCD technology. Furthermore, given the state of the art of photodetector devices, which are also based on CCD technology, the input and output information capacity and rates are nearly matched. Significant improvements in spatial light modulators have been made; future improvements are expected in optical quality, dynamic range, analog control, and cycle rates.

As noted earlier, the performance of a matched filtering system is dependent on the signal scale, orientation, aspect angle, and other parameters. Some researchers sought to find algorithms that might be invariant to scale or orientation, but at the sacrifice of, say, becoming sensitive to signal position. More recently, activity has been focused on ways to sample many test patterns and to construct a filter that has its performance optimized to a composite feature set of the signal. Several authors address these concepts in this issue.

In a broader sense, pattern recognition concepts can also be applied to processing one-dimensional time signals. The ambiguity function and the Wigner distribution are examples of two-dimensional patterns that characterize the time-frequency nature of a signal. Once the signal is reformatted into, say, a Wigner distribution, we may want to classify or recognize the signal source. An example of this at audio frequencies is a voice print; we may wish to automatically recognize what has been said as well as who said it.

This issue of *Optical Engineering* contains eleven papers on the general subject of pattern recognition. In the first paper Yu addresses the question of processing color images, which provides an additional degree of freedom over conventional monochromatic processing systems. In the next paper Gianino and Horner describe some additional work on the use of phase-only filters that improves the light efficiency of the system at the expense of some loss in performance. In applications where laser power is limited, the loss in performance can sometimes be tolerated. In the third paper Psaltis, Paek, and Venkatesh describe the use of a binary, magneto-optic spatial light modulator both as the input device and as the spatial filter.

The next five papers deal with processing algorithms designed to make the performance of the system less sensitive to certain object parameters. Arsenault, Hsu, and Chalasinska-Macukow describe a method for using circular harmonic functions to construct a filter that has rotational invariance while maintaining space invariance. Lahart treats the problem of identifying objects having variable scale. Casasent, Rozzi, and Fetterly describe the use of synthetic discriminant functions in which a training set is used to determine, in some sense, those features of a given object class that can be used to generate filters that are less sensitive to variations in the received signal. Riggins and Butler also use synthetic discriminant functions and show some experimental results using a computer-generated filter. Another feature approach, using the Hotelling trace discriminant criterion, is reported by Gu and Lee. Again, a training set is used to develop the discriminant function so that greater signal variations can be accommodated. These authors show that good performance can be obtained by a single filter, even when the objects in the test set are significantly different.

In the ninth paper Kumar and Carroll show that the Wigner distribution function can be used for pattern recognition and that the performance is similar to that of matched filtering. Since the Wigner distribution is a time-frequency display, it may be possible to implement other operations whereby feature information can be more easily extracted. In the following paper Easton, Ticknor, and Barrett note that computing the Wigner distribution for a two-dimensional object requires a four-dimensional display space. They then show how the Radon transform can be applied to reduce the initial two-dimensional operation to a series of one-dimensional operations on line-integral projections. This information can then be processed to obtain knowledge of the full Wigner distribution.

In the final paper Marks discusses a method for removing the effects induced by multiplicative periodic degradations. Depending on the ratio of the period to the image resolution, the information may be heavily aliased. Marks shows that, under certain conditions, the data can be reconstructed through the use of a second periodic function.

Although progress has been made, pattern recognition remains a difficult problem to be more completely solved, whether the operations are performed optically or digitally. As more tools are developed, both in terms of devices and algorithms, solutions to specific problems will emerge. These solutions will then form the basis for attacking the more difficult general pattern recognition problem. Perhaps someday I will have a robot that will find and retrieve my bags for me.