Object recognition benefits from fast new algorithms

Rohit Patnaik and David Casasent

Advanced correlation filters, combined with the kernel method for pattern analysis, improve classification performance.

General object recognition is a specific application of pattern analysis in which the target object may be present in different orientations. Aspect-view differences and large distortions represent one category (see Figure 1, top row). Objects can also be present at different distances, producing scale differences and images with less detail (see Figure 1, bottom row). Finally, if the location is unknown, a classification algorithm must be applied to all possibilities in the test input. Certain types of classifiers can achieve this with the fast methods currently available, which make them preferable for use in general object recognition.

One category of classifiers is the distortion-invariant filter (DIF). These filters can be implemented efficiently using fast Fourier transforms (FFTs) to locate shifted and multiple objects. A single DIF can handle all aspect-view and a number of scale distortions. DIFs use combinations of training-set images that are representative of the expected distortions in the test set. Many different filters have been developed to address the various distortion problems.1 Our recent work2 demonstrates a new approach that combines DIFs and the kernel technique,3 addresses the need for fast online filter shifts, and improves filter performance.

In kernel-based versions, a classification algorithm is written in terms of vector inner products (VIPs). The VIP sample x and y is written as a kernel function, K:

\[ K(x, y) = \Phi^T(x)\Phi(y). \]

\( \Phi \) is some nonlinear mapping to higher-order space. To use data in that space to take advantage of higher-order correlations, the algorithm must be expressed in terms of VIPs. The kernel function can then be used to evaluate the transformed data, even though \( \Phi \) is unknown. We use this technique with DIFs to form ‘kernel DIFs.’

Prior work4–6 considered several kernel DIFs and applied them to recognition of near-frontal-face images. In this limited application, test faces were registered and centered using the coordinates (assumed known) of several landmark features. Thus, there was no need to apply the filter to shifted versions of the faces. The VIP was calculated for only one location and without consideration of the need for fast online filter shifts. For general object recognition when a shift search is necessary we recently showed7 that kernel versions of the synthetic discriminant function (SDF)1 filter represents the only realistic choice. In addition, we noted that the kernel SDF filter was formulated incorrectly in earlier work, did not use the full higher-order pixel correlations, and could not implement shifts efficiently.

To allow use of fast FFT online implementation of shifted kernel SDF filters, we also demonstrated7 that a polynomial kernel with parameter \( p \) may be used:

\[ K(x, y) = (x^Ty + 1)^p. \]

Similarly acceptable is a Gaussian kernel with parameter \( \sigma \):

\[ K(x, y) = \exp(-||x - y||^2/2\sigma^2). \]

In filter synthesis we must select \( p \) (or \( \sigma \)) and which training-set images to include in the filter. Using training- and validation-set data we automated2 this process.

Kernel SDF filters give 5% improvement on unseen confuser-object-rejection performance and they can handle a 20% larger

Figure 1. For general object recognition an object in a background must be classified in the presence of several distortions, such as aspect-view differences (top) due to different orientations and scale differences (bottom) caused by different ranges.
range of scale than standard DIFs. Likewise, on a difficult discrimination problem involving two very similar objects, these filters delivered noticeably better performance (see Figure 2).

Our work indicates that the number of online calculations for a kernel SDF is directly proportional to the number of training-set images included in the filter. Still more recent results reduce this figure by nearly 50%, which is significant.

For general object recognition in which the location is unknown, classifiers that handle shifts using fast online methods are attractive. We have demonstrated that combining the kernel method and DIFs is promising, but only kernel SDF filters can be efficiently applied for different shifts using fast FFTs and they are, therefore, preferable. Our initial results indicate that, as a consequence of their ability to capture higher-order data correlations, these filters can handle a larger range of distortions than standard DIFs and they are better able to discriminate between two very similar objects. Future work will examine whether kernel SDFs can improve upon performance of standard DIFs when considering recognition in the presence of depression-angle variations.

Support for this work by Raytheon Missile Systems (Tucson, AZ) is gratefully acknowledged.

Author Information

Rohit Patnaik and David Casasent
Department of Electrical and Computer Engineering
Carnegie Mellon University
Pittsburgh, PA

References