Advances in cryoEM and single-particle reconstruction have led to results at increasingly high resolutions. However, to sustain continuing improvements in resolution it will be necessary to increase the number of particles included in performing the reconstructions. Manual selection of particles, even when assisted by computer preselection, is a bottleneck that will become significant as single-particle reconstructions are scaled up to achieve near-atomic resolutions. This review describes various approaches that have been developed to address the problem of automatic particle selection. The principal conclusions that have been drawn from the results so far are: (1) cross-correlation with a reference image (“matched filtering”) is an effective way to identify candidate particles, but it is inherently unable to avoid also selecting false particles; (2) false positives can be eliminated efficiently on the basis of estimates of particle size, density, and texture; (3) successful application of edge detection (or contouring) to particle identification may require improvements over currently available methods; and (4) neural network techniques, while computationally expensive, must also be investigated as a technology for eliminating false particles.

Key Words: electron microscopy; image processing; single-particle reconstruction; object detection; pattern recognition.

1. INTRODUCTION

Cryo-electron microscopy and single-particle image analysis are emerging as powerful techniques in structural biology. In particular, these techniques can be applied to large macromolecular assemblies lacking symmetry, which are difficult to study by other methods. Ever-improving resolution is currently being achieved for a diverse range of macromolecular complexes. Among recent examples, resolutions of 11.5 Å have been reported for the Escherichia coli ribosome (Gabashvilli et al., 2000), a very large, multisubunit particle; 22 Å for bovine complex I (Grigorieff, 1998), an 890-kDa membrane protein; and 21 Å for the catalytic subunit of the DNA-dependent protein kinase (Chiu et al., 1998), a relatively small, 460-kDa soluble protein assembly. Some quite high-resolution results reported recently; e.g., 7.5 Å for the 50S ribosomal subunit (Matadeen et al., 1999) and 9.5 Å for the spliceosomal U1 small nuclear ribonucleoprotein particle (Stark et al., 2001), have used significantly fewer particles than had been required (for the same resolution) in previous work with virus particles, etc. The discrepancy in previous work between claimed resolution and previous work has not yet been reconciled.

As radiation damage imposes strong limitations on the electron doses used to collect micrographs of biological macromolecules, images of individual particles have a low signal-to-noise ratio. The low signal-to-noise ratio can be overcome only by averaging a large number of images. The number of particle images required for a 3D reconstruction increases dramatically with the desired resolution (e.g., 73,000 particles were used in the ribosome reconstruction at 11.5 Å resolution). It is estimated that the number must be increased to about 1 million before it is even physically possible to reach “atomic” resolutions, i.e., better than 4 Å (Henderson, 1995; Glaeser, 1999), when images of the currently available quality are used.

As a first step in image processing, particles from each micrograph are selected either manually, using interactive graphics software as shown in Fig. 1, or by computer-assisted, semiautomated methods. Either way, particle selection becomes a very labor-intensive step in image processing as one works toward ever-increasing resolutions. Thus, automation of particle selection will be necessary to prevent this stage from becoming a serious bottleneck. In light of the growing demand for data sets of ever-increasing size, the time now seems appropriate to review the progress that has been made to date toward the goal of fully automated particle selection.
Several approaches to automatic particle selection have been proposed and have met with varying degrees of success. The approaches that we cover in this review include methods that make use of various forms of template matching, local comparison of intensity values, edge detection, quantitative measures of the local image texture/statistics, and neural networks.

2. TEMPLATE MATCHING

In template matching, an object is detected in an image by evaluating a score for the match between the image and a reference image for each possible object location. Template matching is an important method for object detection and the classical approach is the matched filter/conventional cross-correlation (Abu-Naser et al., 1998; Turin, 1960), discussed in Section 2.1. Object detection occurs when the score exceeds a chosen threshold. A number of methods of template matching have been tried for automatic particle selection (and object detection in general), including cross-correlation-based methods and modified cross-correlation-based methods (including the use of a synthetic discriminant function filter). These various approaches to template matching are described in the following sections.

2.1. Cross-Correlation (Matched Filtering)

Numerous methods have been described for automatic particle selection by cross-correlating a reference image with the micrograph image,
c(x', y') = ∑ₓ ∑ᵧ f(x, y)g(x + x', y + y'), (2.1.1)

where f(x, y) is the image and g(x, y) is the reference. As illustrated in Fig. 2, the operation called for by Eq. (2.1.1) involves shifting the reference relative to the image, multiplying the two, summing the values of the product, and plotting the result at the position (x', y'). In some cases, an azimuthally averaged particle image is used as a reference to avoid the computational cost of examining each possible orientation of the (unaveraged) reference image (Frank and Wagenknecht, 1984; Thuman-Comimike and Chiu, 1995). In other cases the reference image may be something as simple as a two-dimensional Gaussian function the size of which is close to that of the particles that are being sought.

The cross-correlation equation is typically computed using a fast Fourier transform-based algorithm, taking advantage of the correlation theorem. The Fourier transform of the image is multiplied by the complex conjugate of the Fourier transform of the reference and the inverse Fourier transform is computed to obtain the cross-correlation function,

\[ c(x', y') = F^{-1}\{F\{f(x, y)\}F\{g(x, y)\}^*\}, \] (2.1.2)

where F indicates the Fourier transform operation and \( F^{-1} \) the inverse Fourier transform. In effect, \( F\{g(x, y)\}^* \) serves as a filter, \( H(s_x, s_y) \), which modifies the value of \( F\{f(x, y)\} \) before an image is computed again with the inverse Fourier transform. \( H(s_x, s_y) \) is, in fact, known as the matched filter. The idea is that the product \( F\{f(x, y)\}H(s_x, s_y) \), and thus the function \( c(x', y') \), will be maximized when \( F\{f(x, y)\} \) closely matches \( H(s_x, s_y) \), an event that will, of course, happen automatically when \( f(x, y) \) and \( g(x, y) \) are themselves closely similar.

Difficulties are nevertheless encountered with successful peak detection in the cross-correlation function, due to spatial variation of the image intensity and noise in the low-contrast images. Spatial variation and noise in the images result in difficulties in detecting up to 50% of the particles. Problems of this type can be reduced by additional preprocessing of the images, such as bandpass filtering and morphological image processing to identify black areas (Saad et al., 1998). (The topic of morphological image processing is described by Gonzalez and Woods, 1992.) However, the level of false positives remains high and the particles that are detected as candidates must still be screened manually. Another drawback of the cross-correlation function/matched filter, in addition to its sensitivity to spatial variations in intensity, is that it can be sensitive to the type of significant variations that can occur in images of the object that are due to the three possible rotation angles. The letter “A”, for example, does not correlate well with itself when rotated by even a relatively small angle. As a result, to detect multiple views of the same object, it may prove necessary to use multiple references to account for variations in image features with varying particle orientation.

The SDF–MINACE filter represents one approach to detecting multiple views of the same object/particle, and it is described in Section 2.2. The approach described by Ludtke et al. (1999) also combines results from a small number of rotationally averaged references. In their method, the individual cross-correlation maps, each from a different reference (representing one of the more common particle views), are combined by selecting the maximum value at each pixel location from the set of cross-correlation maps.

An alternative method is to ignore the variations that depend upon particle orientation and to simply compute the cross-correlation of the micrograph image with a Gaussian function whose full width at half maximum is equal to the diameter of the object. This “blob correlation” method is in use both as an independent method, without other factors being considered, and as an initial stage for generating data windows that are screened by another method (such as the texture-based method described later in Section 5.2) (Lata et al., 1995). As many as 50% of the particles detected from ribosome micrographs by the blob correlation method (before subsequent screening) in the Lata et al. paper were still false positives.
2.2. Modified Cross-Correlation-Based Methods

Stoschek and Hegerl (1997) applied a modified cross-correlation-based method that uses a synthetic discriminant function (SDF) (Casasent, 1984; Ravichandran and Casasent, 1994) in conjunction with the constraint of minimum noise and correlation energy (MINACE). Casasent (1984) proposed the SDF filter as an improved method for cross-correlation with multiple references (Casasent, 1984; Caulfield and Maloney, 1969), and later work suggested modifications that improve the robustness of the SDF filter to noise (Ravichandran and Casasent, 1994). The inverse filter for a given reference image is given by

$$H(s_x, s_y) = \frac{X^*(s_x, s_y)}{|X(s_x, s_y)|^2}, \quad (2.2.1)$$

where $X$ is the Fourier transform of the reference image (and $X^*$ is its complex conjugate). The product of the inverse filter, $H(s_x, s_y)$, and $X(s_x, s_y)$, fills up Fourier space with unit value. The inverse Fourier transform of this product is a delta function, which is an ideal result in the sense that the correlation peak is sharp. However, the behavior of the inverse filter in Eq. (2.2.1) in the presence of noise is poor. Inspection of Eq. (2.2.1) suggests that the problem arises at those spatial frequencies where $|X(s_x, s_y)|$ is small, i.e., close to the zeroes of the power spectrum of the reference.

The SDF–MINACE filter attempts to obtain the sharp peak produced by the inverse filter while avoiding its instability in the presence of noise. For $N$ reference images $X_i$, ($i = 1, \ldots, N$), an SDF filter has the properties

$$\sum_{k=1}^{K} X_{ik} H_k^* = u_i \quad \text{for} \quad i = 1, \ldots, N, \quad (2.2.2)$$

where $H_k$ denotes a vector element in which the $K$ Fourier coefficients are in lexicographic order, identified by index $k$. Similarly, $X_{ik}$ is a matrix element denoting the $k$th Fourier coefficient of image $i$. Appropriate choices are made for $u_i$ in the equation to reject ($u = 0$) or select ($u = 1$), the corresponding image. Equation (2.2.2) ensures that there will be a value of 1 at the location of a perfect match to one of the selected reference images in the “modified correlation” computed using the SDF filter. However, in the vicinity of the detection peak there may be peaks for false matches (or side lobes) with values larger than 1. To avoid generating side lobes near the detection peak and, also, as the solution for $H_k$ is underdetermined when there is more than one reference image, additional constraints may be applied. Useful constraints on the SDF filter are described by Ravichandran and Casasent (1994).

Stoschek and Hegerl used the SDF–MINACE filter, which is obtained by solving Eq. (2.2.2), subject to the constraint of minimum noise and correlation energy. This constraint is described by the equation

$$\sum_{k=1}^{K} D_k |H_k|^2 \Rightarrow \min. \quad (2.2.3)$$

 Appropriately chosen coefficients, $D_k$, provide a solution to Eqs. (2.2.2) and (2.2.3) for all $N$ and robustness against noise,

$$D_k = \max(S_k, |X_{ik}|^2; i = 1, \ldots, N), \quad (2.2.4)$$

where $S_k$ is the (unknown) noise spectrum. It is common to use a white noise model, with a constant value $S$ for all $S_k$. Stoschek and Hegerl scale $S$ appropriately with respect to the spectra of the reference images to obtain the best performance of the filter; i.e., suitable values of $c$ are used in the equation

$$S = c \cdot \max(|X_{ik}|^2; i = 1, \ldots, N; k = 1, \ldots, K). \quad (2.2.5)$$

In practice, $c$ is a parameter that must be determined empirically for a particular specimen and a given noise level, i.e., by trials of particle detection using the SDF–MINACE filter for different values of $c$.

The method also made use of a circular harmonic expansion—a single circular harmonic component (CHC) from the reference is used to construct the filter to achieve rotational invariance (Hsu et al., 1982), but this is unfortunately done at the expense of the signal-to-noise. In the case of nearly circularly symmetrical objects (such as “flat views” of rings and spherical viruses), taking the zeroth-order CHC is equivalent to the wide use of azimuthally averaged references (Frank and Wagenknecht, 1984; Thuman-Commike and Chiu, 1995). The authors found more difficulty in detecting “side views” in their test data using the second CHC and this may indicate that there is not enough signal-to-noise to detect particles if the target object does not have a large percentage of power in any one CHC. Hsu et al. (1982); Arsenault and Hsu, 1983) discuss the problem of making the best choice of the order of CHC to use in the filter as well as choosing the origin for the...
c(x', y') = \sum_{x} \sum_{y} \left[ f(x + x', y + y') - \bar{f}(x + x', y + y') \right] \left[ g(x, y) - \bar{g} \right] 
= \left\{ \sum_{x} \sum_{y} \left[ f(x + x', y + y') - \bar{f}(x + x', y + y') \right]^2 \right\}^{1/2} \times \left\{ \sum_{x} \sum_{y} [g(x, y) - \bar{g}]^2 \right\}^{1/2} 
\tag{2.3.1}

where \( f(x, y) \) is the image of size \( N \) by \( M \) pixels and \( g(x, y) \) is the reference of size \( N_w \) by \( M_w \) pixels. \( \bar{f}(x, y) \) is a local average, and the summation is carried out over a small area, chosen to be the same size as the reference image, for each point in the micrograph image. The term

\[
\sum_{x} \sum_{y} \left[ f(x + x', y + y') - \bar{f}(x + x', y + y') \right]^2
\]

is proportional to the local variance in the micrograph image, and it can be computed efficiently using one of two algorithms described by van Heel (1992). Successful applications of this approach to identifying areas in noisy images have been reported for IR satellite data by Goshtaby et al. (1984) and for photographic images by Quam and Hannah (1974). The correlation coefficient defined in Eq. (2.3.1) has the desirable properties that its maximum possible value is 1.0, it is automatically compensated for variations in the local (average) intensity, and areas of the image with high variance (due to contaminants, for example) are discriminated against if they do not have a strong correlation to the reference particle.

2.4. Other Related Template-Matching Methods for Object Detection

Other alternatives to improve the performance of template matching include the phase only matched filter (POMF) (Horner and Gianino, 1984; Horner and Leger, 1985; Dickey and Hansche, 1989) and the symmetric phase only matched filter (SPOMF). Cross-correlation supplemented with the use of peak shape information has also been proposed to improve the performance of template matching (Caprari, 1999; Miller and Caprari, 1999). Most of these suggestions remain untested in the context of electron microscopy, although the POMF has been used for phase origin determination in merging of images of two-dimensional crystals (Thomas and Schmid, 1995).

The POMF (as well as its variants) is suggested as being useful for object detection without being sensitive to variation in the image intensity because the phase preserves information about the locations and structures of objects but is independent of the image energy (Oppenheim and Lim, 1981). The POMF is described by

\[
H(s_x, s_y) = \frac{X^*(s_x, s_y)}{|X(s_x, s_y)|},
\tag{2.4.1}
\]

where \( X(s_x, s_y) \) is the Fourier transform of the reference. It can be readily seen that the POMF is closely related to both the matched filter and the inverse filter that were discussed above.

3. EDGE DETECTION

An approach that uses connected components labeling was developed by Harauz and Fong-Lochovsky (1989). Their method has an initial step of edge detection, the effectiveness of which is illustrated in Fig. 3 with an example of images of ribosome particles that are embedded in vitreous ice. Edge detection is then followed by labeling of connected components and symbolic processing, as will be described below.

Component labeling consists of identifying distinct "blobs" or objects in a binary image by assigning labels to adjacent pixels so that pixels that are "connected" receive the same label (Fong-Lochovsky, 1987). Component labeling was carried out using a binary image, obtained by thresholding an image (the "edge image") obtained by edge detection, and connected edge regions (representing the same ob-
ject) receive the same label. This allows edges to be manipulated as symbolic objects. Subsequently, in the symbolic processing stage, closely neighboring components are enclosed with “bounding boxes.”

Symbolic processing (Ballard and Brown, 1982) consists of representing higher level knowledge about the detected objects (such as shape and size) by using symbolic objects and applying predefined rules to deal with the objects. A symbolic object consists of simple, describable properties of the expected object (including its location). The properties and objective rules for handling them can be precisely defined by a human expert.

Harauz and Fong-Lochovsky (1989) used symbolic objects of bounding boxes that are described only by their size and location (and are completely represented by just the coordinates of two opposing corners). Objects that are too close and also large bounding boxes may be rejected in order to exclude aggregates or contaminants. Additional properties may be useful for other specimens, allowing other expert criteria to be applied. The authors used median filtering followed by an edge detector, based on the subclass of linear–median hybrid filters (Heinonen and Neuvo, 1987; Neuvo et al., 1987), to avoid difficulties with edge detection in the presence of noise (Peli and Malah, 1982; Rosenfeld and Kak, 1982). The median filter applied to an image replaces the gray level of each pixel with the median of the gray levels in a neighborhood (e.g., a 5 × 5 pixel square area) of that pixel (rather than the mean, for example). The difficulties encountered with edge detection in noisy EM images nevertheless remain a problem (Lata et al., 1995).

Edge detectors that are less sensitive to noise have been developed in other fields (Canny, 1986; Marr and Hildreth, 1980; Shen and Castan, 1992), so the approach described above may still provide a useful model that an improved method for automatic particle selection might emulate. The edge detection method can be readily applied to quite heterogeneous particles, occurring in a wide variety of views. One reason for the common use of edge detection in machine vision is its insensitivity to variations in illumination conditions (in the context of images less noisy than those commonly encountered in cryoEM) (Davies, 1997). Nevertheless, the method, used by itself, does not have a way of excluding contaminants that may be similar in size to the particles.

4. PARTICLE DETECTION BASED ON INTENSITY COMPARISONS

4.1. The Cross-Point Method of Intensity Comparisons

Boier Martin et al. (1997) described the cross-point method for automatically selecting spherical virus particles. (The method is also limited to objects with uniform internal density in projection. An alternative method, developed by Kivioja et al. (2000) for viruses that appear hollow in projection, is described in the next section.) In overview, it consists of four steps. First, there is initial preprocessing of the image by histogram equalization and subsam-

FIG. 3. An example of the effectiveness of edge detection. (a) A micrograph showing ice-embedded ribosome particles. The micrograph was part of a dataset made available to participants at the “Single Particle Reconstruction from Electron Microscope Images” course held at the Pittsburgh Supercomputing Center on July 21–24, 1999. Comparable data are available as part of the SPIDER software distribution (Frank et al., 1996). (b) An edge image obtained by applying the Shen-Castan edge detector (Shen and Castan, 1992) to the image in (a). The edge image was computed using the Shen program. Source code for the Shen program together with other image processing software is available from Parker (1997).
pling. Particles are then identified by the double-scan procedure, which is described below. Subsequently, in a clustering step, pixels belonging to the same object are identified in a way similar to that used in the algorithm for identifying contiguous edges in the method of Harauz and Fong-Lochovsky (1989). At this stage, clusters that are either too large or too small are then rejected. In the final postprocessing step, checks are carried out for false positives and missed particles (false negatives) that may have been rejected as a part of a large cluster. False hits are identified by comparing the average image intensity inside the particle with that of the background. Missed particles are avoided by thinning the clusters just after the clustering step. Thinning is carried out by removing the outermost pixels from clusters and can effectively disconnect the clusters that have merged into single clusters.

In the double-scan procedure, \( r \) is specified as the radius of the particles to be identified. The image is scanned horizontally, row by row from top to bottom (Fig. 4 describes application of a single scan). Pairs of pixels at distance \( r + 1 \) (in the horizontal direction) are compared and the difference in image intensity is tested against a threshold. If the difference is larger than the threshold, then the lower intensity is tested against a threshold. If the difference in image intensity between a circular area of pixels at distance \( r \) (in the horizontal direction, that are located at a distance \( r + 1 \)).

**4.2. Comparison of Intensities on Rings Rather Than at Points**

Recently, a method (implemented in a program called ETHAN) for detecting virus particles based on the comparison of intensity between a circular area and a surrounding ring was published (Kivioja et al., 2000). Putative particles are further validated by a number of tests, including roundness, proximity of peaks to each other, and peak height. In the procedure, the micrograph is initially filtered using what the authors call a ring filter. If the intensity of the pixel at position \((x, y)\) in the \(N \times M\) image is \(f(x, y)\), the value of the filtered image at position \((x', y')\) is given by

\[
R(x', y') = \frac{\sum_{x=1}^{M} \sum_{y=1}^{N} f(x, y) M_b(x' - x, y' - y)}{\sum_{x=1}^{M} \sum_{y=1}^{N} M_b(x, y)} - \frac{\sum_{x=1}^{M} \sum_{y=1}^{N} f(x, y) M_s(x' - x, y' - y)}{\sum_{x=1}^{M} \sum_{y=1}^{N} M_s(x, y)},
\]

where \(M_s\) and \(M_b\) are mask functions, defined as

\[
M_s(x, y) = \begin{cases} 
1, & \text{if } \sqrt{x^2 + y^2} < r \\
0, & \text{otherwise}
\end{cases}
\]

\[
M_b(x, y) = \begin{cases} 
1, & \text{if } r < \sqrt{x^2 + y^2} < (r + b) \\
0, & \text{otherwise}
\end{cases}
\]

The method essentially involves convoluting the image with a reference consisting of scaled (normalized) versions of a uniform, positive ring and a negative filled circle placed at the ring's center. As this is a centrosymmetric object, the convolution can also
be described as carrying out a cross-correlation with such a reference. This is indeed a good reference for detecting, by cross-correlation, spherical viruses that appear hollow in projection and have the same radius as the ring. The authors describe a real-space implementation of their method, although they mention that implementation of a faster, Fourier-based approach for correlation was left as a subject for future research. The method was found to be highly successful in detecting virus particles. It was found that testing for roundness was valuable for excluding false positives.

5. TEXTURE-BASED METHODS

5.1. Variance

A procedure proposed by van Heel (1982) is based on computing the local variance over a small area, for each point of the image field, as illustrated in Fig. 5. High values of the local variance indicate the presence of an object, since the structure of the object generates internal contrast in addition to that of the shot noise that is present in the background. However, this technique detects all objects, even aggregates and contaminants and, to some extent, noise. As discussed in the next section, we suggest that this and more elaborate methods may be better suited to discriminating whether candidate particles of the correct size, found previously by other methods, should be retained in subsequent steps of the data analysis.

5.2. Linear Discriminant Analysis

A more elaborate method depends on screening an initial set of data windows obtained by convoluting the micrograph image with a Gaussian function. Depending upon one’s point of view, this first step can be described as (1) matched filtering in which the reference object is a Gaussian function simulating an object of the size of the particles being sought or (2) low pass filtering, designed to eliminate the higher frequencies where the signal-to-noise ratio is in any case expected to be poor. This preprocessing step is then followed by a peak search. Pixels in square areas (surrounding peak positions) in the micrograph are copied into small images of candidate particles, or data windows. The data windows identified from the peak search are then screened with a linear discriminant function evaluated from feature vectors computed for each of the data windows.

The feature vectors used by Lata et al. (1995) were obtained by computing nine image parameters—eight were based on the statistics of pixels in the data windows and one was an estimate of particle area. In the training phase, representative data windows are selected by the user to represent one of three categories (particle, noise, or junk), and linear discriminant functions are determined. Training consists of Fischer linear discriminant analysis in which basis vectors are determined with directions giving optimal separation between the different categories in the feature space, in the sense of maximizing between-category variance and minimizing within-category variance. During automatic particle screening, the discriminant functions are used to place new data windows into one of the same three categories.

Linear discriminant analysis is a fairly well-es-
established method for statistical pattern recognition and was described by Duda and Hart (1973) and in more detail by Dillon and Goldstein (1984) in the context of more general applications in statistics.

The method has the advantage that conducting a separate search for each particle view is not needed. As with other approaches that attempt to avoid multireference searches, however, many false positives still get through the discriminant analysis, and manual editing of the data set is still required as a final step.

6. NEURAL NETWORKS

Although neural networks have not yet been applied extensively to the “pattern recognition” problem of particle selection, we briefly review the principles of how they function, to encourage further investigation. The tempting feature of neural networks is the potential that they have to capture, and to take into consideration, a greater degree of complexity than one might normally be aware of when attempting to codify rules for identifying images of single particles.

Artificial neural networks originally drew their inspiration from biological neurons. Artificial neural networks may be used for pattern recognition—both supervised and unsupervised classification. They have been used successfully in various applications, including face recognition (Pentland and Choudhury, 2000), speech recognition (Morgan and Bourland, 1995), and automatic target recognition (Bishop, 1995; Roth, 1990). Neural networks commonly used for pattern recognition include the multilayer perceptron (MLP) (Bishop, 1995; Merelo et al., 1999), learning vector quantization (LVQ) (Kohonen, 1990; Maribini and Carazo, 1994), and the radial basis function neural network (Bishop, 1995). The feedforward neural network, a type of MLP, is briefly described in this review. The interested reader is referred to papers (Kohonen, 1990; Maribini and Carazo, 1994; Merelo et al., 1999) and monographs (Bishop, 1995) for further details and information about other types of neural networks.

The simplest neural network is a perceptron, which is illustrated in Fig. 6a. The output of the perceptron is described by the equation

\[ y = f \left( \sum_{i=1}^{n} (w_i x_i + w_0) \right) . \]  

Its inputs are the components of a feature vector, \( x \). The scalar weights \( w_0, w_1, w_2, \ldots, w_n \) define the performance of the perceptron. \( f(x) \) is called the activation function, and typical choices for \( f(x) \) are nonlinear functions such as the threshold function (with \( f(x) = 0 \) for \( x < 0 \) and \( f(x) = 1 \) for \( x \geq 0 \)) or the logistic sigmoid function, \( 1/(1 + e^{-x}) \). More complicated neural networks consist of a number of nodes (usually single perceptrons) and of interconnections between them. In these more complicated networks, some of the nodes have inputs from other nodes, weighted according to their interconnections.

A commonly used neural network architecture is the feedforward neural network illustrated in Fig. 6b. Feedforward networks can consist of several layers of nodes with interconnections between the nodes of different layers. Typically, however, inter-
connections are allowed only between adjacent layers. In the feedforward neural network, connections back to an earlier layer (loopbacks) are not allowed (although this more complicated architecture is also useful for some applications). A single-layer neural network (i.e., with no hidden layers) uses a hyperplane decision surface in feature space and will be less successful in classifying patterns when the regions for each class in feature space have more complicated boundaries than can be described by sets of hyperplanes (Bishop, 1995); i.e., the boundaries between the classes in the high-dimensional feature space are planes describable by linear equations. Neural networks with one hidden layer are capable of dealing with most classification problems, and they can be used for fitting most functions in regression problems. Neural networks with two hidden layers can learn arbitrary decision surfaces (Lippmann, 1987; Lapedes and Farber, 1988) and can fit any bounded continuous function (Lapedes and Farber, 1988). Neural networks with more hidden layers may be useful in particular applications, e.g., in learning invariances (such as translational or rotational invariance in image recognition).

Neural networks may be used for both supervised and unsupervised classification. In each case, the type of learning rule differs. The learning rule specifies how the weights should be updated to train the neural network. In supervised learning the weights are adjusted so that the outputs match target values for the input pattern vectors in the training set, as measured by the error function. A typical choice is the sum-of-squares error function,

$$E = \frac{1}{2} \sum_{i=1}^{N} \|y(x_i; w) - t_i\|^2 \quad (6.2)$$

where \(y(x_i; w)\) is the output vector for feature vector \(x_i\), \(t_i\) is the target vector, and \(w\) is a vector with the weights for the neural network. In unsupervised learning, the learning rule does not require that the outputs match some targets.

Although neural networks have not previously been applied to automatic particle selection, a series of papers (Maribini and Carazo, 1994; Merelo et al., 1998, 1999; Pascual et al., 2000) has investigated the use of neural networks in electron microscopy for the related tasks of supervised and unsupervised classification. It thus appears that neural networks could be used as one of the tools in automated (rather than manual) elimination of false positives that are contained in an initial set of candidate particles.

Supervised classification is possibly of more interest for automatic particle classification. An example of this application was described in the earlier section on linear discriminant analysis, where linear discriminants were used to place images into one of three categories or classes (two of which are used to reject putative particles). Neural networks may be used in a similar way. LVQ was used for supervised classification of previously aligned and centered images of the helicase T antigen of SV40 (Merelo et al., 1998; San Martin et al., 1997) into type 1 (side views) and type 2 (top views). The data set of 1817 images of size 50 \(\times\) 50 pixels was initially classified by hand. Training of the neural network was carried out using feature vectors obtained by preprocessing the images. The authors preprocessed images by subsampling them to obtain images of smaller dimensions (5 \(\times\) 5). The pixel values in the reduced-size images were used as the components of the feature vectors. This preprocessing step is necessary because the difficulty of determining the parameters or weights of the neural network increases rapidly with the number of dimensions.

Another paper (Merelo et al., 1999) described work on automatic classification of particles using neural networks. The performances of LVQ and MLP neural networks were compared. Also, the advantages of determining the parameters of LVQ or MLP using a genetic algorithm were investigated in a comparison with conventional algorithms. Genetic algorithms (Goldberg, 1989) are a class of algorithms for global optimization that draw their inspiration from natural selection and evolution. In this work, training and classification were carried out using feature vectors computed from particle images that had previously been rotationally and translationally aligned. The feature vectors were obtained by computing rotational power spectra from the particle images. This method of preprocessing has the advantage of reducing the number of parameters that have to be used in the neural network, as described above. Prior to training, 933 particles (SV40 large T antigen) were manually placed into one of five classes—no symmetry or noisy, twofold symmetry, threefold symmetry, fourfold symmetry, and sixfold symmetry. It was found that training MLPs with a genetic algorithm (G-PROP) outperforms the other approaches in terms of successful classification of particles.

### 7. DISCUSSION

As described, a variety of methods (Frank and Wagenknecht, 1984; Thuman-Commike and Chiu, 1995; Stoscheck and Hegerl, 1997; Lata et al., 1995; van Heel, 1982; Harauz and Fong-Lochovsky, 1989; Boer Martin et al., 1997) for automatic particle se-
lection have been proposed in the literature. The majority of these methods have been based on cross-correlation (Frank and Wagenknecht, 1984; Thuman-Commike and Chiu, 1995; Stoscheck and Hegerl, 1997), at least for the selection of images of candidate particles. This selection can be further improved, however, by using the local variance or other statistical properties to discriminate particles from false positives of various types.

A number of the methods described have entered routine use in a semiautomated, computer-assisted mode (Frank et al., 1999; Ludtke et al., 1999). Candidate particles are selected automatically, but a human operator is still required to prune the gallery of putative particle images to reject the images that do not represent true particles. It may also be necessary for the human operator to adjust parameters for each electron micrograph, such as the correlation peak threshold that determines which particles will be selected.

Moving to higher resolutions will involve such a large increase in the number of particles that even this level of human input will have to be eliminated. While it is not an unreasonable task to prune (edit) a gallery of 5000 to 10 000 candidate particle images, this task may nevertheless take about 1 day to complete. Extending such an effort to a few months, as would be required to select a million particles or more, begins to sound like an intolerable assignment.

Important lessons have been learned from the attempts that have been made to develop a procedure for automatic particle identification. These lessons lead us to summarize with the following points:

1. Candidate particles can be identified with reasonable efficiency by use of a matched filter, although this approach becomes computationally intensive when multiple projection views and orientations of particles must be tested.
2. Performance of the matched filter is likely to be improved when template-matching methods that address issues of spatial variation in image intensity are used. The correlation coefficient represents one such approach.
3. Robust edge detection techniques also offer promise as an independent method for particle identification, but their performance on images of the type obtained by cryoEM remains to be tested.
4. The use of additional parameters such as the size of the particle and the statistical properties of image intensities within the area of the particle helps to discriminate true positives from false positives.
5. Further tools are still needed, however, to reduce the fraction of false positives to such a low level that a final stage of human “pruning” and supervision can be eliminated.

We thank Joachim Frank for permission to use the ribosome test images for the examples shown in Figs. 3 and 5; comparable data are available as part of the SPIDER software distribution (Frank et al., 1996). This work was supported in part by PACI Subaward 776 from the University of Illinois for NSF Cooperative Agreement No. ACI-9619019 and by the Director, Office of Science, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

REFERENCES

Gabashvilli, I. S., Agrawal, R. K., Spahn, C. M. T., Grassucci,
AUTOMATIC PARTICLE DETECTION


