

Detection and Quantification of Protein Spots by *Pinnacle*

Jeffrey S. Morris and Howard B. Gutstein

Summary:

Accurate spot detection and quantification is a challenging task that must be done effectively in order to properly extract the proteomic information from two-dimensional (2-D) gel electrophoresis images. In Morris, et al. (2008), we introduced *Pinnacle*, an automatic, fast, effective non-commercial package for spot detection and quantification for 2-D gel images and subsequently have developed a freely available gui-based interface for applying the method to a set of gels. In this chapter, we overview *Pinnacle*, and in a step-by-step manner describe how to use the software to obtain spot lists and quantifications to use for comparative proteomic analysis.

Key Words: Automated methods, Image processing, Spot detection, Wavelet denoising

1. Introduction

A key bottleneck in proteomic research involving 2-D gel electrophoresis has been the lack of automatic, effective, and reproducible methods for analyzing the gel images. It could be argued that these analytical problems have been one of the key factors inhibiting this technology from reaching its potential in biomedical research.

Given that the proteomic information in the gel images is contained in spots resulting from migrating proteins on the gel, the critical preprocessing steps involve detecting, matching, and quantifying spots across gels. Sensitive spot detection algorithms are essential, since any spots not detected in this step are lost to subsequent analysis, leading to lost discoveries. Accurate spot matching is crucial, for mismatched spots result in different proteins put together in follow-up analyses, compromising the ability to detect which ones are associated with outcomes of interest. Accurate spot quantification is also important in order to minimize measurement error that can reduce power for detecting significant differences.

Until recently, the only available analytic methods were commercial software packages with problematic underlying algorithms and workflows that broke down for studies involving more than a very small number of gels (Gutstein et al. 2008, Morris et al. 2008). These packages followed a traditional workflow that involves detecting spots, determining spot boundaries, and estimating spot volumes on each individual gel, and then matching the detected spots across gels. This workflow has numerous weaknesses. It results in missing data for any spots not detected on all gels, causing analysis challenges and introducing potential bias. It is subject to spot matching errors since cognate spots will be at slightly different locations on each gel, and the matching step is only done after spot detection. Mismatched spots compromise the ability to detect differentially expressed proteins since different proteins are mistakenly grouped together for the analysis. Spot boundaries are also difficult to estimate properly, especially in regions with overlapping spots caused by co-migrating proteins, increasing the imprecision of spot quantifications. This problem is compounded when different spot boundaries are estimated for each gel. Most software packages provide manual editing capabilities to correct this plethora of errors, but this is a time-consuming process that compromises the objectivity and reproducibility of the analysis. Exacerbating the problem, all of these errors worsened as the number of gels in the study increased. This suggested a perverse incentive for researchers to conduct studies with fewer

samples, which can lead to insufficient power to detect significant differences, especially when properly accounting for the multiple testing problem inherent to the simultaneous analysis of hundreds or thousands of proteins. The dominance of these expensive, yet ineffective software packages was partly fueled by the lack of freely available methods for spot detection and quantification, which was especially problematic for researchers performing infrequent 2-D gel studies whose laboratories could not afford the software.

As a solution to this problem, we developed *Pinnacle*, a novel analytical method (Morris et al. 2008) for protein detection and quantification that utilizes a completely different workflow to avoid these problems. Working on pre-aligned gels, *Pinnacle* performs spot detection by searching for pinnacles (local maxima in both dimensions) on a wavelet-denoised (pixelwise) average gel. In the average gel, signals from spots found on many gels are reinforced while white noise is attenuated. Thus, detection based on the average gel can have greater sensitivity than detection done on individual gels, especially for faint spots present on many gels (Morris et al. 2005). In stark contrast to previous methods, spot detection accuracy improves as the number of gels increases. Specificity is further improved by the wavelet denoising that eliminates spurious noise without strongly attenuating the true signal. After spots are detected on the average gel, quantifications are obtained for each spot for each gel by taking the maximum intensity within a defined neighborhood around the detected “pinnacle” (peak intensity value for each spot), while performing local background correction and normalization using one of several approaches. *Pinnacle* is quick and automatic, results in no missing data, and the use of pinnacle intensities instead of spot volumes for quantification obviates the need to detect spot boundaries resulting in greater precision and reliability. In fact, *Pinnacle* enables one to detect potentially co-migrating proteins (or protein isoforms) if multiple peak values are detected in a region of the gel that other spot-based methods would tend to call a single spot.

Using dilution series experiments, we compared *Pinnacle* with two commercially available processing packages that at the time were state-of-the-art, and found that *Pinnacle* resulted in significantly improved validity (higher R^2 with true quantification) and reliability (lower %CV on replicate gels) than the commercial packages, and in a comparative analysis found more significantly different spots across groups (Morris et al. 2008). After our initial development of *Pinnacle*, various software companies developed alternative workflows that worked on pre-aligned gels to minimize problems from spot matching after detection, and improved spot quantification by defining common spot outlines to use across all gels to quantify spot volumes for each cognate spot. Using dilution series, we compared with the first of these methods, *Progenesis SameSpots* (Nonlinear Dynamics), and found that its performance was significantly improved over *Progenesis PG240*, the previous version of Nonlinear Dynamics’ software in terms of spot detection and quantification (Morris et al. 2010). In these comparisons, *Pinnacle* still outperformed *SameSpots*, leading to more spots detected, spot quantifications with higher validity and reliability, and detecting more significantly different spots in a comparative proteomics analysis (Morris et al. 2010). The type of algorithm underlying *SameSpots* is still prevalent in commercial software today.

In this document, we describe the *Pinnacle* and associated graphical user interface in more detail, providing a step-by-step explanation of how to input gel images into the software, perform spot detection and quantification using *Pinnacle*, and then outputting the spot quantification lists for subsequent analysis.

2. Materials

The *Pinnacle* software is a standalone Windows package with a graphical user interface, so a Windows-based system is needed to use the software. Here we mention the required format of the gel images for use with the software.

1. The software requires the gel images be stored as 16-bit .tif files. All gel images for a common analysis should be stored in a common directory.
2. **VERY IMPORTANT:** *Pinnacle* is designed to work with gel images that have already been registered or warped to each other so that the spots are reasonably well aligned on the separate gel images. There is no such registration package included with *Pinnacle*, so the user must use another package to do the registration and warping, save the warped/aligned images as 16-bit .tif files, and then use these for the *Pinnacle* analysis. Various commercial and non-commercial methods are available for performing this registration; some are automated algorithms while others require manual intervention. Some of these methods are described in this book. We use RAIN (Dowsey et al. 2008), which is overviewed in Chapter 18 of this book.
3. Do **NOT** edit the .tif images in Photoshop or any other imaging processing software, as this could alter the quantitative information in the gel image and affect accuracy.
4. When scanning the gels, avoid saturation of the gel images – *Pinnacle* works best when gels have little to no saturation.

3. Methods

To frame our presentation, following is a mathematical representation of the model underlying *Pinnacle*. Let $Y_i(x, y)$ represent the registered scanned gel image $i = 1, \dots, N$, at horizontal location x and vertical location y in the image, with the image sampled on a grid of size $J \times K$. The set of images are assumed to have been registered to each other to establish a matching correspondence in the (x, y) for the separate gels.

$$Y_i(x_j, y_k) = \underbrace{B_i(x_j, y_k)}_{\text{Background}} + \underbrace{N_i}_{\text{Normalization Factor}} \underbrace{S_i(x_j, y_k)}_{\text{Protein Signal}} + \underbrace{e_{ijk}}_{\text{Noise}}$$

The quantity $B_i(x, y)$ represents background artifact for gel i , assumed to follow a smooth spatial stochastic process, e_{ijk} represents white noise, assumed to be independent and identically distributed mean zero-Gaussians with variance σ^2 , and N_i represents a constant multiplicative normalization factor for gel image i , included to adjust for systematic differences in total protein quantity or scanning settings across gels. Our goal in preprocessing is to extract the protein signal $S_i(x, y)$, which is expected to be made up of a convolution of P protein spots centered at locations (x_s^*, y_s^*) , $s = 1, \dots, P$. This can be written as

$$S_i(x, y) = \sum_{s=1}^P Q_{is} \phi_s(x, y),$$

where $\phi_s(x, y)$ is a canonical spot template for spot s , which often can be well-approximated by a 2-D Gaussian density function centered at (x_s^*, y_s^*) , and Q_{is} represents the quantification for spot s from gel i . We will refer back to this equation and notation throughout our presentation.

3.1 Loading gel images into *Pinnacle*, loading and saving project files.

1. To start a new analysis, in the *File* menu, choose *Start New Project*, which will bring up a dialog box to select the directory in which the registered gel .tif files reside.
2. Select the folder containing the .tif files, and then by default the software will load all .tif files in this folder into the program. At this time, two windows open by default, an *Individual Gel* window and *Average Gel* window. Section 3.2 will discuss these windows. Figure 1 contains a screen shot of the *File* menu and these two windows, tiled vertically.
3. If one wishes to only analyze only a subset of the .tif files or if the directory contains other .tif files that are not registered gel image files, then one can choose *Select Gel Images* in the *Action* menu, and mark the set of .tif files corresponding to the desired gels for analysis. At this time, the average gel will update to only include selected gel images.
4. At any time, one can save the current project in the *File* menu using *Save Project* or *Save Project As*, and can reload that project by choosing *Open Existing Project* in the *File* menu. The project file saves all current settings and results.

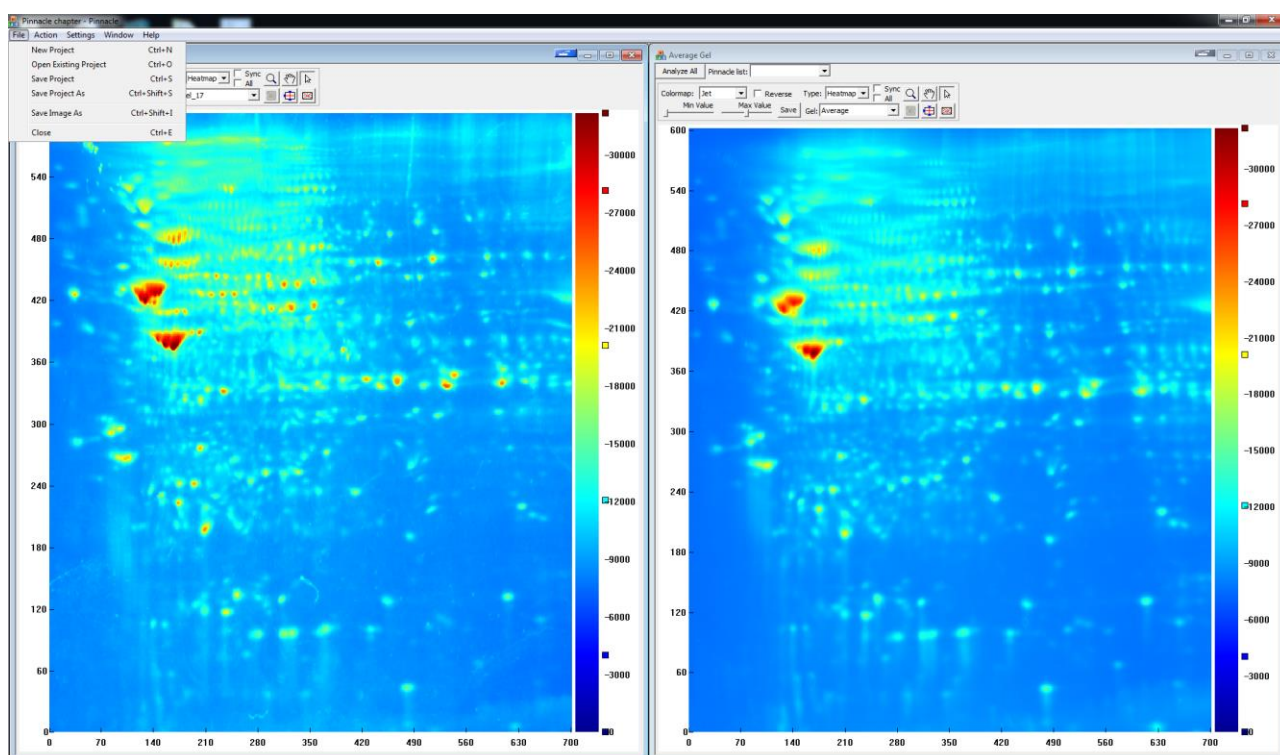


Figure 1: Screen shot showing *File* menu and the *Individual Gel* and *Average Gel* windows.

3.2 Viewing individual gels or average gel

By default, the software opens two graphical windows, one displaying an individual gel and the other displaying the average gel for the entire gel set, computed as the pixel-wise average across all registered gel images. These windows are useful for exploring the properties of your gel images, and have considerable flexibility for choosing which gels and which statistical summary gels are computed and for zooming in and out, moving around the gel, and changing the contrast to allow the exploration of faint and more intense spots. The gel windows can be synchronized together so that any changes/zooming/moving done in one gel window will be mirrored in the other gel windows. After pinnacles are detected, they are indicated on these windows by “x” marks. These gel image displays are useful for exploration of the pinnacle detection performance and to provide a framework for rapid, intuitive hand-editing of the results. In this section we will describe some of the settings and capabilities of graphical windows.

1. By default the *Individual Gel* window plots a heatmap of the final gel in the set that was loaded. If you want to see other gels, select a different one in the *Gel* pop-up list. When the *Gel* pop-up list is selected, turning the mouse scroll wheel will step through the different individual gels. Multiple individual gels can be opened in separate windows by selecting *Open Individual Gels* in the *Window* menu, or all individual gels can be opened using *Open All Individual Gels* (although be careful about doing this if the gel set is very large!). The windows can be arranged in various ways using the *Arrange Windows* in the *Window* menu.
2. **Average Gel:** By default the Average Gel window plots a heatmap of the arithmetic, pixel-wise mean gel. Alternatively, the pixel-wise maximum gel or quantile gel can be computed under the *Gel* pop-up list. The maximum gel contains the maximum intensity for each pixel across all selected gels, while the quantile gel contains a specified quantile (e.g. 0.50 for median, 0.95 for 95th percentile).
3. **Colormaps:** Various colormaps are available for plotting the heatmaps of the gels. By default, *Jet* is chosen, which plots the intensities ranging from colder colors to hotter colors (blue→red). This is our personal choice, as the rich contrast of colors appears to make visual inspection of the spots more intuitive. Some may prefer the images to look more like the physical gels, in which case the *Grey* colormap can be chosen, with intensities ranging from darker to lighter colors (black→white). For any colormap, by checking the box *Reverse*, one can reverse any colormap, e.g. to let white be lower intensities and black be higher intensities for the *Grey* map.
4. **Image Contrast:** The contrast of any heatmap can be adjusted by changing the minimum and maximum values of the colormap, which can be done by the sliders marked *Min Value* and *Max Value*. This is especially useful in investigating the spot detection for fainter or overlapping spots. By decreasing the *Max Value*, fainter spots become more clear and one can see whether they look like “real” spots that should be selected. However, this will tend to make the more intense spots run together. To visually delineate nearby intense spots, increasing the *Max Value* will help one assess whether there are multiple local spots in a particular intense region.
5. **Zooming and Navigating:** By default, the entire gel is shown in the window, but one can zoom and move around the image to investigate local regions of the gel. One can zoom in on the image by selecting the magnifying glass button, and then left-clicking and dragging the mouse up and to the right. Zooming out can be done by left-clicking and dragging the mouse down and to the left. When zoomed, one can move around the image by selecting the hand button and left-clicking and dragging around. The image can be centered on a particular spot by selecting the button with a box and red cross in the bottom middle of the menu bar and then left-clicking on the desired centering point for the gel image.
6. **Synchronizing Gel Images:** When changing the colormaps or contrasts of the images or while zooming or moving around the image, one can synchronize these changes across all gel image windows by checking the box *Sync All*. If one has multiple images opened but only wants to synchronize over a subset, they can check the *Sync* box for each image window to be synchronized. This is useful in exploring the pinnacle detection to see whether the spots in the average gel appear to match well with visual spots in individual gels as a check on the alignment.
7. **3D Plots:** Instead of heatmaps, the gel images can be displayed as 3d plots by selecting *3D Plot* in the *Type* pop-up list. These 3D plots have the same colormap specified by the heatmap. The perspective can be changed by selecting the rotation button in the menu bar and left-clicking and dragging the mouse around.
8. **Saving Gel Images:** One can save the currently active gel image by selecting *Save Image As* under the *File* menu. Supported formats include .jpg, .tif, .pdf, and .ps.

3.3 Spot Detection

Overview: *Pinnacle* performs spot detection on the average gel, given by

$$\bar{Y}(x_j, y_k) = N^{-1} \sum_{i=1}^N Y_i(x_j, y_k) = \bar{B}(x_j, y_k) + \bar{S}(x_j, y_k) + \bar{e}_{jk}$$

where $\bar{S}(x, y) = N^{-1} \sum_i N_i S_i(x, y)$ is the average signal, $\bar{B}(x, y)$ is the average baseline, and $\bar{e}_{jk} \sim N(0, \sigma^2/N)$ is the noise on the average gel, reduced by a factor of \sqrt{N} relative to the noise level in an individual gel because of the averaging. This will increase the signal-to-noise ratio for spots present on at least a certain proportion of all gels, and thus can lead to increased sensitivity and specificity for spot detection. Further, this means that the accuracy of the spot detection should increase with the number of gels in the analysis. These concepts were illustrated by a rigorous simulation study in the context of peak detection for 1D mass spectrometry (Morris et al. 2005), and the same concept applies to spot detection here. **IMPORTANT:** Note that *Pinnacle* is designed to perform spot detection on the average gel, and we don't recommend its application to individual gels.

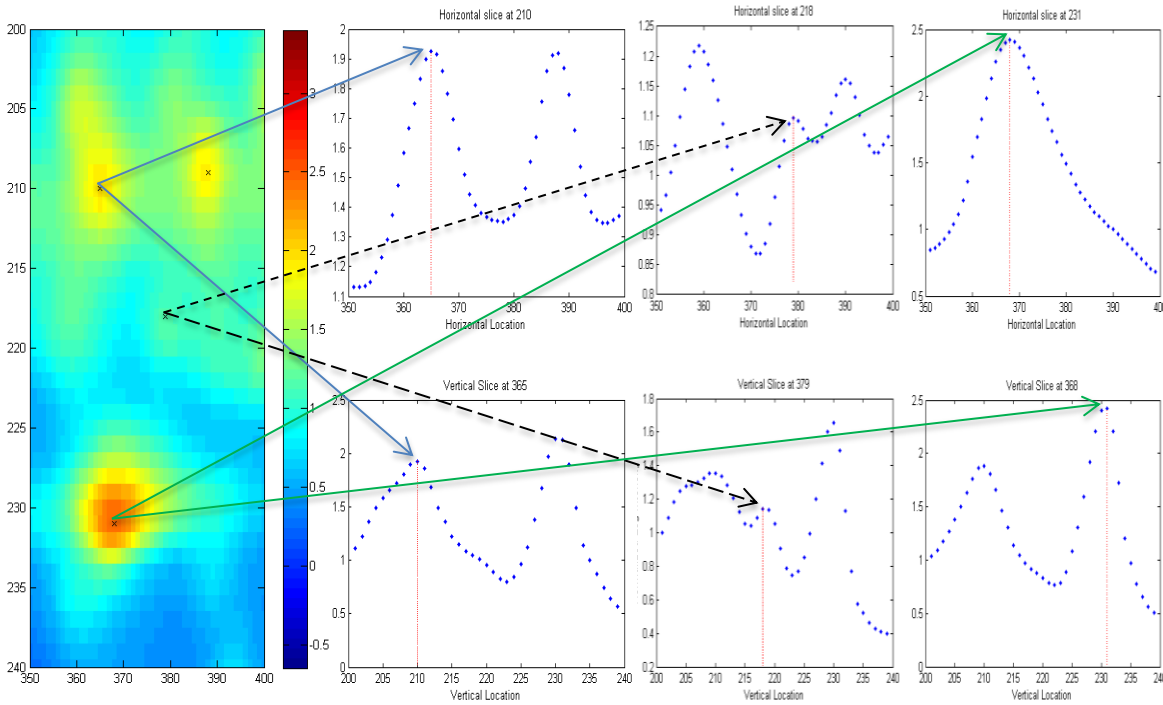


Figure 2: Illustration of a pinnacle. The heat maps on the left show a region of an average gel image, and “x”s mark detected pinnacles. The right three columns illustrate three of the pinnacles in this field, with the pixel intensities for a single row and column corresponding to one of the marked pinnacles plotted in the top and bottom, respectively, with a red line indicating the pinnacle location, defined as a location that is a local maximum in both the horizontal and vertical directions.

Next, wavelet denoising is done to estimate and remove noise (\bar{e}_{jk} in the equation above), and then positions (x_s^*, y_s^*) are identified that are local maxima in both the horizontal and vertical directions (see Figure 2), and for which $\bar{Y}(x, y) > q_\alpha$, where q_α is a **minimum pinnacle size threshold**, defined as the α quantile among all intensities on the average gel $\bar{Y}(x, y)$. Next, we aggregate any pinnacles (x_s^*, y_s^*) and $(x_{s'}, y_{s'})$ that are very close to each other, i.e. $|x_s^* - x_{s'}^*| < \delta$

and $|y_s^* - y_{s'}^*| < \delta$, where δ is a prespecified **neighborhood size**, by keeping spot s if $\bar{Y}(x_s^*, y_s^*) > \bar{Y}(x_{s'}^*, y_{s'}^*)$, keeping spot s' otherwise. While our studies show that *Pinnacle* tends to work very well without any manual editing, the software allows manual editing of the pinnacles, adding or removing them, using a dynamic display including the mean gels and individual gels while allowing zooming in on particular regions of the images, moving around the image, and changing the contrast of the heatmaps to accentuate fainter spots or to separate merged spots. The strategy of detecting on the mean gel along with the graphical interface make manual editing of the spots an efficient and straightforward process.

1. **Defining Image Region:** Before spot detection, one can first define an *image region* S including the regions of the gel where spot detection is desired. The reason for this is to give the user the option to exclude regions of the gels containing obvious artifacts, which are especially common near the boundaries of the rectangular image. The region can be marked either using a *rectangular* or *lasso* tool under *Specify Image Region* in the *Action* menu. All spot detection, background correction, and normalization will be done only using the data in this region.
2. **Detecting, Saving, and Loading Pinnacle Sets:** One can detect pinnacles by choosing *Detect Pinnacles* within the *Action* menu. The pinnacles will be detected using the settings indicated in the *Pinnacle Settings* dialog box in the *Settings* menu, which will be discussed below. One can save the pinnacle set by choosing *Save Pinnacle List* in the *Action* menu. One can load in previously saved pinnacle sets through *Load Pinnacle List*, and can choose among any of the loaded pinnacle sets in the *Pinnacle List* drop-down menu in the *Average Gel* window, which contains a list of all loaded pinnacle lists for the current project. One can open a window containing the pinnacle list by selecting *Pinnacle Window* within the *Window* menu. This window displays (x_s^*, y_s^*) for each pinnacle, the intensity in the average gel, and an indicator of whether it was automatically detected or manually added (see 4 below).
3. **Pinnacle detection settings:** There are a number of different settings that affect pinnacle detection that can be adjusted in the *Pinnacle Settings* dialog box in the *Settings* menu. There are default values for each setting. Figure 3 contains a screen shot of the pinnacle settings window, the pinnacle window, and individual and average gel windows with each detected pinnacle indicated by a red or green “x”.
 - a. **Wavelet denoising settings:** Wavelet denoising on the average gel is used to remove spurious “wiggles” that can manifest as false positive pinnacles. It works by transforming the image to the wavelet domain using a 2-D Daubechies wavelet transform with *wavelet filter length*, thresholding any wavelet coefficients below a specified *wavelet threshold*, and then performing the inverse wavelet transform to construct the denoised image. Both the *wavelet filter length* and *wavelet threshold* can be set in *Pinnacle Settings*. The *wavelet filter length* must be an even number (default 6). Results tend to be robust to the filter length choice. The default value for the *wavelet threshold* is 3.6. Making it larger will result in more denoising but if too large, may smooth away true features, while making it smaller will result in less denoising.
 - b. **Other pinnacle settings:** Signal detection also depends on two other parameters: the *minimum pinnacle size threshold* q_α and the neighborhood size δ , described above. These parameters are set in the *Pinnacle Settings* dialog box. In the *Minimum Pinnacle Size Threshold* box, one specifies α , the percentile indicating the minimum intensity for a pinnacle. Increasing this threshold reduces the number of spurious spots but making

it too large may remove faint but true spots, while making this threshold too small can result in more spurious spots. The *neighborhood size* box indicates the minimum distance between two pinnacles. Any pinnacles closer than this amount are merged together, taking the maximum value. This allows for slight misalignment of gels, since slightly misaligned gels may lead to multiple local maxima in the average gel corresponding to the same cognate protein. Making this parameter too large will merge together different proteins.

- 4. Hand editing detected pinnacles:** We have found that the automatic spot detection through Pinnacle typically performs quite well, especially if one tweaks the minimum pinnacle size threshold or wavelet threshold settings for a particular set of gels, assessing the performance by visually inspecting the average gel and individual gel images. However, the software also provides the capability to hand-edit the pinnacle lists. This is typically done by zooming in on the average gel, moving around, and varying the contrast (min value, max value) to inspect the detected pinnacle, and then removing any that do not look like real spots (e.g. artifacts), or to add pinnacles for any regions that appear to be spots in the gel image window but were not detected. Pinnacles can be added by holding *Ctrl* and left-clicking on a location of either the average gel window or any individual gel window. Each pinnacle is represented in the gel image windows by a red or green “x”, whichever has the best contrast with the gel image at that location. Pinnacles can be removed by holding *Ctrl* and right-clicking while moving the cursor around in one of the gel windows. The *Eraser Size* in the *Pinnacle Settings* dialog box gives the size of the square eraser used to remove pinnacles, effectively “erasing” them from the gel image. Following are some notes and tips:
- It also frequently helps to check *sync all* in the gel image window so one can see both the average gel and individual gels in the regions being investigated.
 - When deciding whether a given pinnacle corresponds to a true spot or not, we typically toggle through the individual gels to see whether we believe it is a real spot or not. This also gives insight into whether the gels are aligned well enough.
 - We typically will zoom in to a range of about 1/5 of the gel rows and columns and systematically pan around the image, adding or removing spots.
 - We have found that this process typically takes on the order of 10-15 minutes for the entire gel set. Generally, less than 10% of the spots that were automatically detected are modified manually.

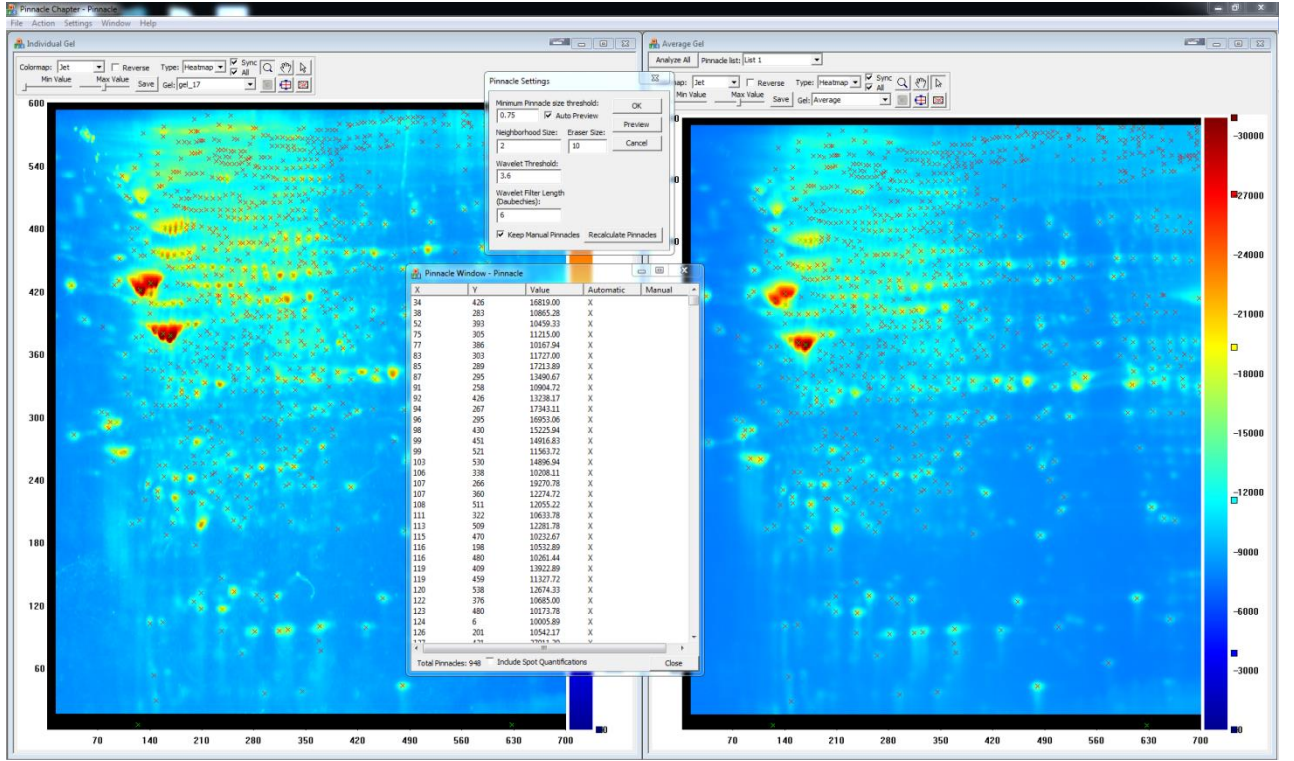


Figure 3: Screen shot showing *Pinnacle Settings* dialog box and the *Pinnacle Window*. Note that the pinnacles are evident as red “x”s in the gel image windows.

3.4 Spot quantification

Overview: Once spot detection is completed, and a list of detected spots and corresponding pinnacles (x_s^*, y_s^*) , $s = 1, \dots, P$ found on the average gel, then we compute relative spot quantifications Q_{is} for each spot and each gel. Rather than estimating the spot boundaries and computing the spot volumes, quantification is based on the pinnacle intensity. We have shown that as long as a cognate spot s has a common canonical spot shape $\phi_s(x, y)$ across gel images, a reasonable assumption in our investigation, the maximum pinnacle intensity and the integrated spot volumes are measuring essentially the same quantity (Morris, et al. 2008). The benefits of using pinnacles for spot quantification are that they are faster and do not require estimation of spot boundaries, potentially reducing measurement error and increasing precision of spot quantifications. As discussed in the introduction, the use of common spot boundaries and registered gel images has improved the current software package performance, but in our view there is still a benefit to quantifying protein amounts by pinnacle intensities rather than integrated spot volumes, since no spot boundaries must be estimated.

Spot quantification is done by first computing the maximum intensity on the individual gel image in a neighborhood around the pinnacle location (x_s^*, y_s^*) on the average gel, i.e. $Y_{is}^* = \max\{Y_i(x, y), x \in (x_s^* - \delta, x_s^* + \delta), y \in (y_s^* - \delta, y_s^* + \delta)\}$, where δ is the specified **neighborhood size** (specified in the *Pinnacle Settings* window). Background correction and normalization are also done to estimate the background $B_{is} = B_i(x_s^*, y_s^*)$ and normalization factor N_i , and then the quantification for spot s for gel i is given by $Q_{is} = \frac{Y_{is}^* - B_{is}}{N_i}$. Optionally, wavelet denoising can be done to estimate and remove the white noise $e_{is} = e_i(x_s^*, y_s^*)$. Note that this algorithm provides well-defined spot quantifications for each cognate spot for each gel, so there is no missing data.

1. **Quantifying Spots:** One can obtain the spot quantifications by selecting *Quantify Spots* from the *Action* menu. Spot quantification will be done according to the *Spot Quantification Settings* in the Settings menu (see Figure 4), as discussed below. In the gel image windows, once the spots have been quantified the red/green “x” marking the pinnacles is recolored black/white, and the “x” in the individual gel image shows the actual location used for quantification for that gel image (see Figure 4). After performing the quantification, the user is prompted to provide a prefix name for the .xls files containing the quantification information. Three .xls files are created, one *_spotsettings.xls containing the settings used for pinnacle detection, *_spotlist.xls containing for each cognate spot the pinnacle location x_s^*, y_s^* , an indicator of whether the pinnacle was automatically or manually added, and the intensity in the average gel $\bar{Y}(x_s^*, y_s^*)$, and *_spotquants.xls containing for each cognate spot and gel the spot number, an indicator of whether automatically or manually detected, the gel label, the gel-specific pinnacle location x_{is}^*, y_{is}^* , and the spot quantification Q_{is} . The quantifications can also be shown in the *Pinnacle Window* under the *Window* menu by checking the box *Include Spot Quantifications* in the *Pinnacle Window* (see Figure 4). At any time, the spot .xls files can be exported using *Export Spot Quantifications* in the *Action* menu.

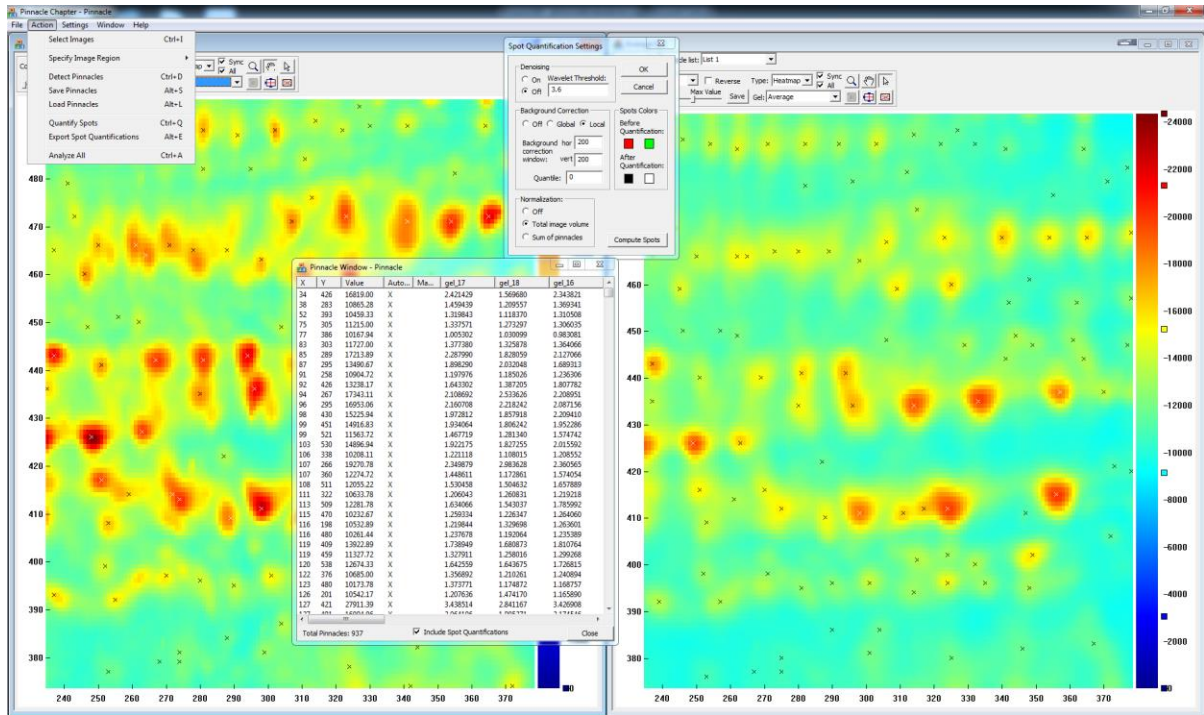


Figure 4: Screen shot showing *Spot Quantification Settings* dialog box and the *Pinnacle Window* with spot quantifications for a zoomed in region of the gel image.

2. **Spot Quantification settings:** There are a number of different settings that affect spot quantification that can be adjusted in the *Spot Quantification Settings* dialog box in the Settings menu (see Figure 4). There are default values for each setting.
 - a. **Denoising:** Denoising can be turned on in order to estimate and remove the white noise errors in the image e_{ijk} . If denoising is done, then the wavelet threshold can be specified (default 3.6). By default, denoising is turned off.
 - b. **Background Correction:** Background correction, if turned on, will subtract an estimated baseline value from each pinnacle to estimate and remove $B_{is} = B_i(x_s^*, y_s^*)$. If *local* is selected (default), then B_{is} is estimated to be either the minimum or specified

Quantile within the rectangle $[x \in (x_s^* - \gamma_x, x_s^* + \gamma_x), y \in (y_s^* - \gamma_y, y_s^* + \gamma_y)]$, where γ_x and γ_y are the specified **horizontal Background correction window** size and **vertical Background correction window** size, respectively. The local minimum is computed when **Quantile**=0. When *global* is selected, the global quantile or minimum within the specified image region for the entire gel is used as the background for all spots.

- c. **Normalization:** Normalization is done to adjust for differences in the total protein content or total image intensity in the individual gel images. The constant N_i is estimated by taking either the **total image volume**, defined as the sum of the gel intensities in the specified image region S , or the **sum of the detected pinnacles**. We use the total image volume for normalization by default.

3.5 Analyze All

Under the *Action* menu, *Analyze All* will do a complete, automatic pinnacle analyses using the current settings, detecting, quantifying and saving all pinnacles while performing background correction, normalization and denoising.

4. References

Clark BN and Gutstein HB (2008). The myth of automated, high-throughput two-dimensional gel electrophoresis. *Proteomics* 8: 1197-1203.

Dowsey AW, Dunn MJ, and Yang GZ (2008). Automated image alignment for 2-D gel electrophoresis in a high-throughput proteomics pipeline. *Bioinformatics* 24: 950-957.

Morris JS, Clark BN, and Gutstein HB (2008). Pinnacle: A fast, automatic and accurate method for detecting and quantifying protein spots in 2-dimensional gel electrophoresis data. *Bioinformatics* 24: 529-536.

Morris JS, Clark BN, Wei W and Gutstein HB (2010). Evaluating the performance of new approaches to spot quantification and differential expression in 2-dimensional gel electrophoresis studies. *Journal of Proteome Research* 9(1): 595-604.

Morris JS, Coombes KR, Koomen J, Baggerly KA and Kobayashi R (2005). Feature extraction and quantification for mass spectrometry in biomedical applications using the mean spectrum. *Bioinformatics* 21, 1764-1775.

Morris JS and Gutstein H, inventors: *Method and Computer-Program Product for Detecting and Quantifying Protein Spots*, Patent Number US 8,031,925, Issue Date 10/4/2011.

5. Acknowledgements

This work was supported by grants from the National Cancer Institute (CA107304, CA160736) and from the National Institute on Drug Abuse (DA18310) and the National Institute on Alcohol Abuse and Alcoholism (AA16157), and these algorithms and software are to subject of Patent Number US 8,031,925 (Morris and Gutstein 2011).

