Framework for Parsing, Visualizing and Scoring Tissue Microarray Images

Andrew Rabinovich, Stan Krajewski, Maryla Krajewska, Ahmed Shabaik, Stephen M. Hewitt,
Serge Belongie, *Member, IEEE*, John C. Reed, and Jeffrey H. Price, *Associate Member, IEEE*

Abstract—Increasingly automated techniques for arraying, immunostaining, and imaging tissue sections led us to design software for convenient management, display, and scoring. Demand for molecular marker data derived *in situ* from tissue has driven histology informatics automation to the point where one can envision the computer, rather than the microscope, as the primary viewing platform for histopathological scoring and diagnoses. Tissue microarrays (TMAs), with hundreds or even thousands of patients' tissue sections on each slide, were the first step in this wave of automation. Via TMAs, increasingly rapid identification of the molecular patterns of cancer that define distinct clinical outcome groups among patients has become possible. TMAs have moved the bottleneck of acquiring molecular pattern information away from sampling and processing the tissues to the tasks of scoring and results analyses. The need to read large numbers of new slides, primarily for research purposes, is driving continuing advances in commercially available automated microscopy instruments that already do or soon will automatically image hundreds of slides per day. We reviewed strategies for acquiring, collating, and storing histological images with the goal of streamlining subsequent data analyses. As a result of this work, we report an implementation of software for automated preprocessing, organization, storage, and display of high resolution composite TMA images.

Index Terms—Automated tissue microarray (TMA) scoring, densitometry/fluorometry, image acquisition, texture segmentation, tissue microarrays (TMAs).

I. INTRODUCTION

TISSUE MICROARRAYS (TMAs) enable more efficient analyses of tissues for cancer tissue expression studies and many other applications, but are laborious to read. Hun-

A. Rabinovich and S. Belongie are with the Department of Computer Science and Engineering, University of California at San Diego, La Jolla, CA 92093 USA (e-mail: amrabino@ucsd.edu; sjb@cs.ucsd.edu).

S. Krajewski, M. Krajewska, and J. C. Reed are with the Burnham Institute for Medical Research, La Jolla, CA 92037 USA (e-mail: stan@burnham.org; maryla@burnham.org; jreed@burnham.org).

A. Shabaik is with the Department of Pathology University of California at San Diego, La Jolla, CA 92093 USA (e-mail: ashabaik@ucsd.edu).

S. M. Hewitt is with the Tissue Array Research Program, Laboratory of Pathology, Centers for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20894 USA (e-mail: hewitts@mail.nih.gov).

J. H. Price is with the Burnham Institute for Medical Research, La Jolla, CA 92037 USA, the Department of Bioengineering, Whitaker Institute of Biomedical Engineering, University of California at San Diego, La Jolla, CA 92093 USA, and Vala Sciences, La Jolla, CA 92037 USA (e-mail: jhprice10@yahoo.com).

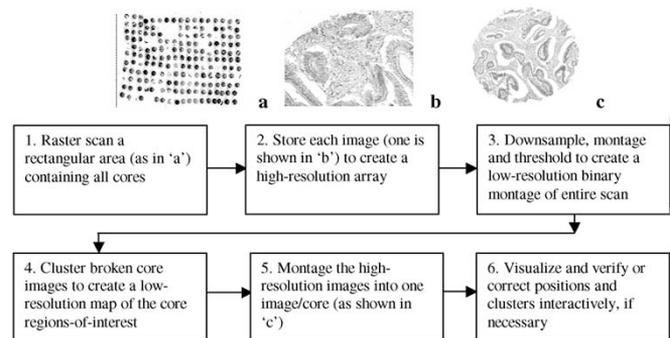Digital Object Identifier 10.1109/TITB.2005.855544

Fig. 1. A flow diagram is presented of the automated acquisition and visualization demonstrated on a colon cancer tissue microarray. The only inputs required are the scan area (x, y, dx, dy) and the number of cores. After these steps are completed, the images are ready for diagnosis/scoring. The image in 'b' is a single field of view from a 20 × objective and 'c' is a montage of images acquired at 20 ×.

dreds to thousands of cylindrical cores (of 0.6–2.0 mm diameter and up to 30 mm long) are biopsied from tissue donor blocks, and arrayed in a single new recipient paraffin block [1], [2]. The new block is sectioned and placed on slides [Fig. 1(a)], [1]. The presence of many tissue samples on one slide dramatically increases the processing and analysis rates. Many patients with a given type of cancer can be studied for protein expression patterns simultaneously by immunohistochemical methods. Tissue microarray blocks can also be analyzed for gene profiles by *in situ* hybridization (ISH) or with fluorescence (FISH) [2]. Researchers thus identify cancer biomarkers which correlate with differences among the arrayed samples, such as histological grade, pathological stage, or patient survival.

The scale of the task depends in part on the number of cores required for each tumor. The number of cores required may be increased by the need to sample more than one region of the tumor to obtain sufficient tissue to represent the entire tumor. When, in fact, one or two cores may be needed to represent the entire tumor in cases of breast carcinoma as indicated by Zhang *et al.* [3] and Camp *et al.* [4], respectively; other tumors known for their heterogeneity like prostate cancer may require more cores [5]. The distribution of the marker to be studied may also impact the number of samples needed from each patient. For instance, some markers such as HER2/neu or ER may have sufficiently low inter-tumor variability to justify the use of one or two cores, whereas more heterogeneous proliferation markers such as Ki-67 would likely require more. It seems reasonable that the number of tissue samples needed may vary as a function of the particular marker and the type of cancer, and requiring more cores increases the analytical challenge in large archival

TABLE I
SUMMARY OF AVAILABLE TMA SCANNERS

System	Optics	Map Cores	Image(s) Displayed	File Formats
A	Bright	Grid	Abstracted	Proprietary
B	Bright/Dark/Fluor	Spot Find	Individual	JPEG/TIFF
C	Bright	Spot Find	Individual	JPEG/TIFF
D	Bright	Grid	Abstracted	Exportable
E	Bright/Fluor	Spot Find	Both	TIFF
F	Bright/ Fluor (Confocal)	Grid	Individual	JPEG/TIFF
G	Bright	?	Abstracted	?
H	Bright/Fluor	?	Abstracted	JPEG/TIFF
Ideal	All	High-Dry Auto	Magnification Range, User-Selectable	JPEG/TIFF

Definitions

A-H are example commercially available systems with the names removed.
Bright: brightfield, or transmitted light microscopy
Dark: darkfield microscopy
Fluor: fluorescence microscopy
Confocal: confocal microscopy in fluorescence
Grid: interactively defined grid to locate each core in the array
Spot Find: the computer finds the locations of the cores automatically
High Dry Auto: the computer automatically locates the cores after scanning with a high-dry objective (0.95 NA 40-100x)
Abstracted: images are displayed by selecting a point on a low resolution image of the entire scan area
Individual: one image is acquired at the center of each core and displayed
Both: the array is scanned at low magnification for locating the cores followed by with a higher power scan that either acquires a single image at the center of each core or creates a montage of the whole core from several images
Magnification Range, User-Selectable: the images are scanned at high magnification and a low, medium or high magnification view (or continuous zoom) of any core is available by random access (upon user demand)
Proprietary: image format unique to the system
JPEG/TIFF: standard format such as JPEG or TIFF image files exportable to most imaging software packages

cohorts [3]. Improving automation promises to simplify increasing the number of samples for greater statistical significance.

Manual analysis of the immunohistochemical results is the primary bottleneck in TMA studies. In our hands, TMA analysis typically takes ∼ 10 times longer than TMA production. Pathologists typically read the intensity or percentage of immunopositive cells in each tissue core to identify cancer-associated changes in gene expression and to make correlations with tumor histology or other variables. This labor-intensive bottleneck makes automated analyses a compelling goal.

The ultimate TMA analysis automation goal is to load a cassette of, say, 100 slides and return the next day to find molecular label measurements in a convenient database, linked to the images of the TMAs. Approaches toward this goal vary. Slide loaders for microscopes; e.g., from Vision BioSystems, have been available for many years. Several complete systems are specifically designed for automated TMA analysis. Manufacturers include Aperio, Biogenex Laboratories, Bacus Laboratories, Universal Imaging, and Applied Imaging. The characteristics of example automatic scanning systems A-H are summarized in Table I to aid description of the predominant approaches. TMAs are most commonly stained with color or fluorescent dyes and less often with immunogold, autoradiographic or quantum dot preparations. To accommodate all techniques, the system would require optics for brightfield (Bright), fluorescence (Fluor), and darkfield (Dark) or reflection (for silver or gold particles). Reflection via a combination of cross-polarized filters and a beam-splitter in place of the usual epifluorescent filters (e.g., the Nikon IGS cube) is preferred by many over darkfield for silver and gold metallic particles.

Different techniques are used to map the locations of the cores on the slide for image acquisition. The user may be required to interactively define the grid points (Grid) prior to scanning. In some systems, the cores are located automatically by image analysis (Spot Find), usually during a low magnification (e.g., 1–4 × or perhaps with a conventional digital camera and lens) scan and thresholding step that is followed by a higher magnification (e.g., 10 ×) scan. Smaller fragments of broken cores can be missed during low magnification scanning and then excluded from the higher magnification rescan, which may result in the loss of diagnostically useful information. The interactive (Grid) technique can add 20–60 min to each scan for TMAs with 500 or more cores. During scanning, one *Individual* image may be collected from each core (i.e., near the center at higher magnification) or the images may be *Abstracted* from a montage of several images that comprise each core. The *Image Analysis* capabilities of the system might then range from simple display to more sophisticated densitometry, fluorometry and/or particle counting (for gold or silver). The best image formats are those that are in common use such as JPEG, PNG, TIFF and BMP. The use of lossless vs. lossy compression should be evaluated to ensure that diagnostic accuracy is preserved. The software should store the image and data files for access by other programs or provide an export function in a common image file format (*Exportable*). Poor image quality can frustrate diagnostic efficiency and accuracy and image quality may be degraded in some systems by a camera that undersamples the optical resolution (i.e., with lower camera than microscope resolution). The ideal system would scan slides stained by any method without human interaction and provide access to low, medium and high resolution images of any portion of a tissue array or core (*User Selectable*). Pathologists sometimes need high magnification images to make diagnoses so the ideal system should be capable of scanning the images at *high-dry* resolution (0.95 NA at 40–100 ×) at high speed (e.g., 6–8 min at 20 × 0.5 NA and 25–30 min at 40 × 0.95 NA). The combination of large computer storage requirements and long scanning times may limit use of high magnification images, but rapid advances in storage and scanning may solve these problems in a few years.

The automated TMA scanning strategy of imaging only the tissue cores after mapping the locations of the tissue cores (*Spot Find*) from a low resolution image adds the most efficiency for slower incremental-motion stage scanning. As speed is increased and systems using continuous-motion scanning become more common, pre-mapped tissue core locations will have less impact on speed except perhaps by better defining the outer boundary of the entire scan area.[1] For this reason, we designed

---

[1]Limitations in the flatness of the specimen and the optics make it difficult to use ever-larger format CCD cameras to increase speed. At high resolution (i.e., NA = 0.95), the specimen can be uneven enough to cause blur within the field of view of the objective, although this may not be as much of a problem with thick tissue sections as with cell monolayers. Very large format cameras would require larger fields of view than present in commercially available microscope objectives. Since microscope objectives are the highest performance manufactured optics available, customized larger field-of-view high resolution (i.e., NA > 0.6) lenses rapidly become prohibitively costly. Lens arrays have been proposed to overcome this problem, but at high resolution (narrow depth of field) this approach would require independent autofocus for each element for typical microscope slides, which can violate flatness by tens of micrometers.

our software to automatically collate tissue core sub-images from an exhaustive scan of the entire TMA area on a slide. Here we report software for automatically locating, montaging, organizing, storing and displaying high resolution composite TMA images. Our goals to achieve "ideal" operation (see Table I) included: 1) use of brightfield and fluorescence (or reflection) optics; 2) fully automated scanning up to high-dry resolution (no low magnification pre-scan or interactive grid finding required); 3) random access to low, medium, and high resolution images upon user request; 4) convenient data entry for diagnoses and scoring; and 5) common image formats and data storage for easy import/export to/from other databases [12].

II. METHODS

A. Tissue Sections Construction of Microarrays

Tissue microarrays were constructed as previously reported [13]. The use of human tissue was reviewed and approved by the UCSD and Burnham Institute Institutional Review Boards (IRBs). Experimental animal tissue was harvested using IRB-approved protocols. A brief summary of TMA preparation follows.

1) Tissue Preparation: Normal tissues for immunohistochemical analysis were derived either from human biopsy and autopsy material (Department of Pathology, UC San Diego) or from adult mice of various strains. The tissues were fixed in either neutral-buffered 8%–10% formalin, Z-fix zinc-buffered formalin solution (Anatech Ltd, Battle Creek, MI) [14], B-5 [15], or Bouin's solution (Sigma, Inc. St. Louis, MO), and embedded in paraffin.

2) Tissue Array Construction: Tissue microarrays contained 130 arrayed specimens represented by 0.6 mm or 1 mm diameter cylindrical cores acquired from paraffin blocks of human or mouse tissues, cut into 4–5 μm thick sections. Tumor TMAs were constructed by removing 2–5 cylinders of 1 mm diameter tissue from representative areas of each archival paraffin block and arraying them into a new recipient paraffin block with a manual tissue arrayer (MTA-1, Beecher Instruments, Silver Spring, MD). Serial sections (4 μm) were applied to 3-aminopropyltriethoxysilane-coated slides (Sigma), as described by Rentrop *et al.* [16] The prostate tissue microarray was constructed as described previously by Kononen [1].

3) Immunohistochemistry and Fluorescence: Deparaffinized tissue sections were exposed to polyclonal antibodies (PAB) generated against synthetic peptides and confirmed to be specific for Bid, a Bcl-2 family protein, which like its other family members, is correlated with apoptosis, or programmed cell death [13]. The sections were immunostained using a diaminobenzidine (DAB)-based detection method as described in detail, employing either an avidin-biotin complex reagent (Vector Laboratories) or the Envision-Plus-Horse Radish Peroxidase (HRP) system (DAKOCytomation) using an automated immunostainer (Dako Universal Staining System) [17], [18]. The dilution of antiserum was 1 : 8000 (v/v). The immunostaining procedure was performed in parallel using preimmune serum to verify specificity of the results. Initial confirmations of antibody specificity also included experiments in which antiserum was preadsorbed with 5–10 μg/ml of either synthetic peptide immunogen or recombinant protein immunogen. For some preparations, samples were stained with hematoxylin-eosin. Alternatively, tissue sections were deparafinized, stained with DAPI and imaged with a 20×0.5 NA Nikon objective. The DAPI solution (Molecular Probes, Eugene, OR) consisted of 75 ng/mL DAPI, 10 mM Tris, 10 mM EDTA, 100 mM NaCl, and 2% 2-mercaptoethanol. Tissues sections were incubated in DAPI solution for 45 min prior to sealing the slides with coverslips.

B. Microscopy Instrumentation and Image Acquisition

High throughput microscopy (HTM) instrumentation that operates with both dry and oil-immersion objectives has been described previously, is commercially available from Q3DM/Beckman Coulter as the IC-100, and is summarized here. We made minor modifications to our IC-100 equivalent high throughput microscopy system, which was originally designed for automated fluorescence image cytometry [6]–[11], and used it to automatically scan TMAs. Previous reports demonstrated scanning speed of up to 3 fields/s [8]. The current instrument is based on autofocus that works in 0.25 s with a precision (SD) of <100 nm [8], and lamp stabilization with 10 to 30 times better CVs than the conventional Hg vapor lamp [9]. The autofocus circuit [8] was modified for brightfield tissue section scanning by adding a larger integrating capacitor, 1100 nF, to handle the increased brightfield sharpness signal from hematoxylin-stained tissue. A monochrome CCD camera (Cohu, San Diego, CA, model 6612-3000, 8- or 10-bit digital output with 659×494 9.9×9.9 μm^2 pixels) captured the images and the software cropped them to 640×480. A Nikon Eclipse 300 microscope was used with 20×0.50 NA Plan Fluor Ph1 DLL and 40×0.95 NA Nikon Plan Apo objectives. The areas magnified to each pixel in the CCD camera (calibrated using a Nikon micrometer with 10 μm per line) were 0.538×0.538 μm^2/pixel and 0.270×0.270 μm^2/pixel, respectively, for fields of view of 344×258 μm^2 and 173.0×129.7 μm^2, respectively. Although some of these are phase contrast objectives, a brightfield condenser was used. Theoretically, the phase ring in the objective alters the frequency spectrum of the image by performing optical Fourier filtering, but this effect is expected to be subtle and we could not discern any differences in the images obtained using phase objectives with a brightfield condenser. This instrumentation operates robustly on large scan areas in experiments on cell-by-cell analysis of 5 million cytocentrifuged cells per slide that found as few as 1:20 million cells in ultra-rare cell detection [10].

1) Scanning the Slide and Storing the Images (Steps 1 and 2): Slide scanning and data storage were carried out as follows. The slides were placed on the HTM system, the operator entered the positions of the corners of the scan area, focused manually at the first field of view and initiated the scan of step 1 in Fig. 1. The system autofocused on every field and remained at or near focus even over areas with no tissue [8]. For the microarrays, about 6,000 RGB images were acquired automatically (and stored as step 2 of Fig. 1) by scanning in a raster pattern Cohu CCD
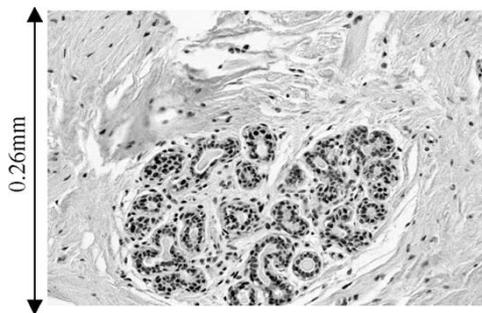
Fig. 2. One field of view is shown from a scan of a normal breast tissue core stained with hematoxylin and DAB. This is 1 of 6,000 images that compose the scan of the whole TMA area with a $20\times$ objective.
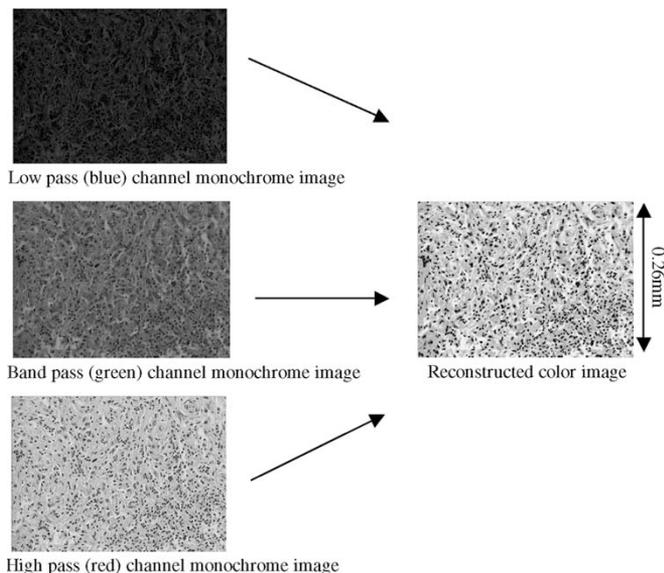

Fig. 3. A reconstruction of a color image from three monochrome channels is shown ($20 \times$ objective).
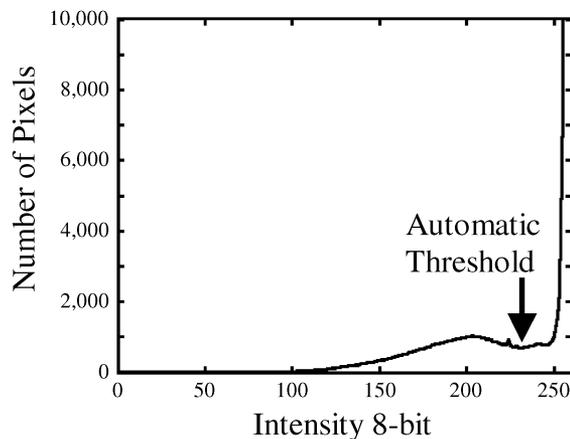

Fig. 4. An intensity histogram is shown of the reduced-resolution montage of the entire tissue microarray. The background and foreground were separated via a local minimum in a specified intensity range.

camera mounted via a $1 \times$ relay lens and color filters in an automated filter wheel. An example image is shown in Fig. 2. The DAB and HRP immunohistochemistry stains cover most of the visible spectrum. Red-green-blue (RGB) images were collected using a low pass filter with 50% transmittance at 505 nm, a band pass filter with a maximum transmittance at 517 nm and 50% transmittance at 500 nm and 570 nm, and a high pass filter with 50% transmittance at 605 nm, respectively. Images were stored as 8-bit/color, 24-bit images, and reconstructed as illustrated in Fig. 3. The images of each core were then processed, reassembled and placed in a database, described as follows.

C. Image Processing and Analysis

Automated scanning and image storage steps 1 and 2 of Fig. 1 were carried out by the modified HTM system as described previously. Our Tissue Chips software, written in Matlab 6.1, then processed the raw images and loaded them into a database for viewing and scoring as shown steps 3–6. For reassembly, the raw images were reduced in resolution by $10 : 1$ (down-sampled by linearly-interpolated averaging of 3.16×3.16 neighborhoods of pixels into one pixel) and binary images were produced by automatic intensity thresholding (step 3) to locate the fields of view representing tissue. Sectioning sometimes breaks the cores into groups of pieces. The resolution reduction of $10 : 1$ was chosen to so that even small fragments of tissue were retained. In step 4, the fragments of tissue were clustered automatically into a single grouped data structure for each core (as described further below). In step 5, the grouped data structures were used to control reconstruction of a montaged image for each core. The final montaged images of each core were reconstructed and stored at the native high resolution of the original scan. The intervening blank images were then deleted. Finally, in step 6, thumbnail images (resolution-reduced by 100:1) of each core were displayed in a regular rectangular grid for confirmation of the locations.

1) Locating the Tissue Sections in the Images (Step 3): The tissues were located by the following image segmentation steps. The RGB images were reduced in resolution by 10:1 to speed automatic intensity thresholding. Histograms of the three reduced-resolution RGB channels were averaged together to create a single grayscale 8-bit intensity profile. The RGB-average histogram was smoothed using a Savitsky–Golay filter (available in Matlab) [21]. Savitsky–Golay filters are convolution filters of length j that exactly replicate fitting a polynomial of degree k and derivative order l over a given data set (we used $j = 7, k = 5$, and $l = 0$). There are several methods for automatic intensity thresholding for image segmentation including those by Ramesh *et al.* [22], Otsu [23], Duda and Hart [24] and Gonzalez and Wintz [25]. We chose an iterative quadratic fit (Newton's method) [26] of the smoothed, averaged histogram to find the local minimum between tissue section pixels and background pixels (see Fig. 4). The upper and lower boundaries of the quadratic fit range were the two largest peaks in the histogram (foreground and background). The lower intensity tissue histogram peak was sometimes broad and generated too low of a threshold. To compensate, the minimum threshold was
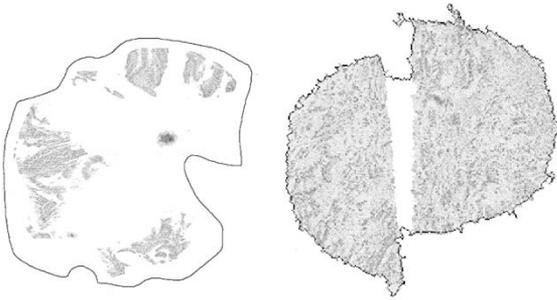
Fig. 5. The grouping of two fragmented tissue cores (acquired from colon cancers sections with a $20\times$ objective) in the microarray using K-Means clustering in Cartesian coordinates is shown. The black line around each group encircles the corresponding automatically combined area. The example on the left is an extreme case of core fragmentation during TMA preparation (the apparent border region is caused by debris and small pieces of tissue that are not visible at the resolution shown here).

set to 183 empirically. Intensities above the automatic threshold were labeled background and intensities equal to or below the threshold were labeled tissue.

2) Clustering Broken Cores (Step 4): In sectioning the paraffin block, tissue cores often break apart and must be appropriately grouped to preserve the accuracy of the data. The K-means clustering algorithm [27] provides an automatic method for grouping objects based on distance. K-means determined the centroids and boundaries of the individual cores after it was initialized with the pixel coordinates of the thresholded objects and K, the total number of expected cores on the slide. Thus, the K-means algorithm clustered broken pieces of cores into single database objects. The grouping of the cores is based on the Euclidean distances $D(p, c)$ in Cartesian coordinates. The cost function applied was

$$D(p, c) = \sqrt{(x_p - x_c)^2 + (y_p - y_c)^2}$$

where x and y are the image coordinates, p is a given pixel, and c is a given cluster center. Examples of broken cores that were clustered are shown in Fig. 5. K-means clustering regroups broken pieces nearest each other.

Choosing K in the K-means algorithm is an instance of the model order selection problem. Although many automated techniques have been proposed, model order selection is still an open problem in machine learning. Methods such as K-nearest neighbor (KNN), hierarchical clustering (HC), and mean shift (MS) may also be used for grouping broken cores, however, these approaches do not resolve the model order selection problem: in KNN the number of nearest neighbors must be specified, in HC the granularity of clustering must be chosen, and in MS the radius of the hyper-sphere must be selected [34]. In our application we have chosen K-means since the number of cores is known during TMA construction.

3) Montaging the Raw Images (Step 5): Thresholding and clustering provide computer representations of the locations and pixel maps of each core from the reduced-resolution image of the entire scan area; each core is uniquely identified and numbered. These image segmentation maps were then used to reconstruct high-resolution montage images of each core from

the raw high resolution images in the array of raster-scanned fields-of-view acquired during the scan. The resulting montaged images (super-bitmaps) were then stored for later use and the raw scans containing intervening blank images were deleted. These super-bitmaps of each core may be stored without compression in BMP format or with lossless compressed with PNG or TIFF (using LZW) formats. Any core in the resulting array of montaged super-bitmaps is then made available by the software for viewing at zoomed-in native resolution by clicking on thumbnail images reduced in resolution by $100 : 1$.

4) Interactive Correction of Misalignment and Clustering (Step 6): Mistakes in automatic clustering have not been observed, but are possible. For example, if a core is broken into distant pieces that lie close to another core, the automatic assignment may be incorrect. Thus, the errors are dependent in part on the proximity of the cores and the distance between fragments. If cores are completely missing, initialization of the K-means clustering algorithm may be incorrect (a potential issue with a slide loader where differences in core number from slide to slide might not have been entered manually) and the remaining cores may not be placed in the proper array location. These errors can also be controlled by TMA preparation. In step 6, the user can interactively correct the positions of the cores in the grid and any mistakes in automatic clustering. Our experience thus far indicates that these kinds of errors are rare enough to be corrected by user interaction. If errors in K are more common that expected, other approaches for automatic detection and correction can be considered.

5) Entering and Displaying Data: A set of $100 : 1$ reduced-resolution, correctly positioned array of thumbnail sketch core images was stored in a database with each core directly linked to the montaged, native resolution images as described above. This enabled creation of a convenient system for viewing and scoring the images at multiple magnifications, at the discretion of the operator. Our Tissue Chips software creates a numbered display of thumbnail images of the cores through which a series of higher magnification views and options for data entry are available by mouse clicks. Tissue Chips includes information input and retrieval options, including tumor histological grade, patient information, diagnosis and treatment, as well as the position of a given core in the array (e.g., for making corrections as described above). Users can populate the database in two ways: by manually typing the information into the windows shown, or by importing the data from existing databases in standard formats (e.g., tdf, cdf). The fine cellular detail that is sometimes important for diagnostics can be displayed quickly and conveniently as needed at resolutions as high as the raw acquisition.

III. EXPERIMENTAL RESULTS

Fully automated histological image acquisition was tested firstly on conventional tissue sections stained with hematoxylin and eosin (brightfield) and DAPI (fluorescence), and secondly on TMAs stained with hematoxylin and DAB. Figs. 6 and 7 show example images from large tissue sections, and Fig. 8 shows examples from TMAs. Fig. 6(left) shows a 100:1 reduced-resolution montage of 9×9 fields of view (of a $1.12 \times 1.49 \, \text{mm}^2$
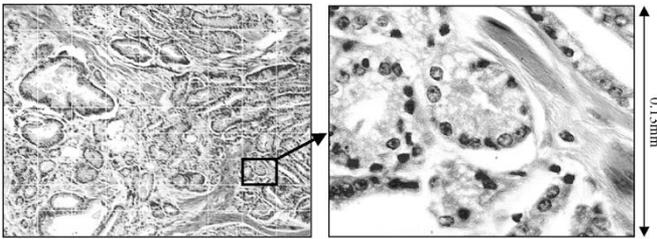
Fig. 6. Left: A downsampled (100:1) montage is shown of 9×9 fields of view (grid superimposed) of a 1.12×1.49 mm^2 area acquired with a 40×0.95 NA Nikon Plan Apo objective. The specimen presented is a section of hematoxylin and eosin stained prostate adenocarcinoma. Right: the selected field of view is shown at native resolution.
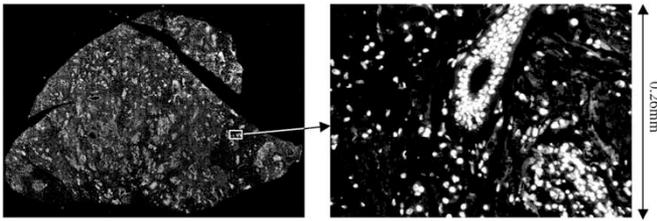

Fig. 7. Breast cancer tissue stained with DAPI is shown. Left: A montage (51×51 fields of view) of a 1.27×1.69 cm^2 area containing a tissue section scanned with a 20×0.5 NA Nikon objective is shown. Right: One field of view is shown at native resolution.

area) from the middle of a prostate adenocarcinoma tissue section scanned with a 40×0.95 NA objective. The right-hand portion of Fig. 6 shows a selected field of view at native resolution. The depth of field of a 0.95 NA objective is 0.55 μm ($d = \lambda/\text{NA}$ [2], [28] $\lambda = 500$ nm), which is much smaller than the 5-μm tissue section. The autofocus algorithm finds the best combined focus for the entire field of view as tested through several focus positions. This drives it to select the axial center of the section. As can be seen in the native resolution image (Fig. 6, right), sharply focused tissue components are visible throughout the field of view. Most of the nuclei appear in focus, with only a few probably lying outside this plane of focus. All of the fields of view in this scan demonstrated high focus quality. Precise autofocus in the axial center of the tissue section at each field can produce sharply focused images even where section thicknesses vary.

The ability to scan fluorescently stained samples was tested on large breast cancer tissue sections stained with the DNA-binding fluorochrome DAPI. The 51×51 field-of-view montage (2×3 cm^2 area) shown in Fig. 7 (left) was scanned using a 20×0.5 NA objective (left). A selected field of view is shown on the right at native resolution. The images were all sharply focused and the brightly stained cell nuclei are visible along with dimly stained cytoplasmic and extracellular matrix staining. Note that the section was broken into two larger pieces and several very small fragments around the periphery. It is common to place more than one tissue section on each slide in clinical practice. The automatic clustering algorithm groups the fragmented tissues into the number of tissue sections supplied by the operator. Thus, images of conventional slides, with more than one tissue section per slide can also be archived by section when multiple

fragmented sections are placed on the same slide and whether stained for fluorescence or brightfield.

Tissue microarrays represent a much larger number of tissue sections on a single slide to be organized, archived and scored. Fig. 8 shows the low magnification (100 : 1 reduced-resolution, upper left), medium magnification (middle left) and high magnification (lower left) images of a TMA with 162 cores (9×18). The low magnification images can also be displayed in a regular rectangular array pattern as shown at the upper right. The user can select individual cores for display at medium and high resolution for scoring and data entry (lower right). Images can also be displayed at medium resolution in sequence for scoring. High resolution views are available with a mouse click at any time if detailed display is needed for diagnosis or scoring.

Automatic acquisition and collation of images for scoring provides a platform for easily creating, accessing and manipulating the information in a database. When the scores and images are linked to the clinical information about each core, the data entry and retrieval shown on the right in Fig. 8 can be generalized for searching databases for expression patterns and predictors. Complete clinical information (including treatment and outcome data) can be stored on other databases along with references to the TMA images and scores for advanced prognostic studies that protect patient confidentiality.

The Tissue Chips software developed to extract, montage, and archive images and provide a digital interface for data entry can be applied to raw images collected on any system. To demonstrate this, a TMA was sent to Aperio Technologies Inc. (Vista, CA, www.aperio.com) for scanning. Raw images were extracted through the Aperio software interface and processed by Tissue Chips to produce the 100:1 reduced-resolution image in Fig. 8 (top right, upper). The automatically reconstructed tissue cores are also shown arrayed in Fig. 8 (top right, lower) for comparison with the raw reduced-resolution array. This array of extracted cores was aligned in Fig. 8 by using rectangles of equal size instead of the variable sized ones at the upper left of Fig. 8.

IV. DISCUSSION

This report demonstrates methods for automated preprocessing, extraction, reconstruction, display, and archiving of images derived from tissue sections and TMAs. Both our own slide scanning method and that of Aperio's involves scanning the entire area containing the tissue sections at the prescribed final resolution. While in continuous-scanning systems little or no time is lost imaging blank areas because the system is already moving to the next tissue core, imaging the blank areas increases scanning time in our IC-100 compatible system. An alternative automated method is to scan at low magnification/resolution to first obtain a map of the tissue sections and then reposition for imaging each tissue section (e.g., in instruments by Zeiss, Universal Imaging Corporation, Biogenex Laboratories, Bacus Laboratories, and Applied Imaging Corporation). In automated fluorescence TMA analysis reported by Camp *et al*. 2002, a 512×512 μm^2 inscribed image (1024×1024 pixel images for 0.5×0.5 μm^2 sampling and 1.0 μm optical resolution that undersampled a
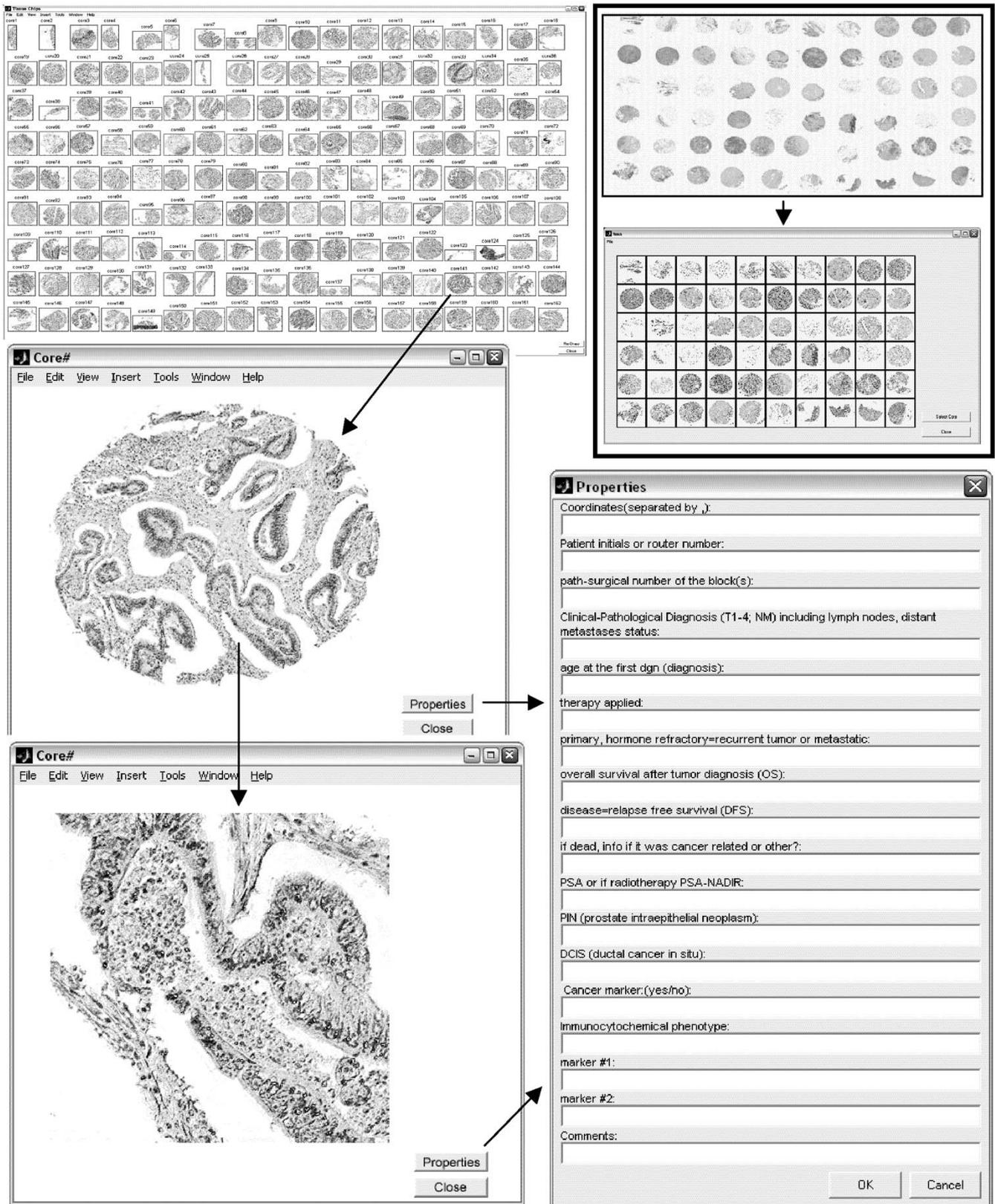
Fig. 8. Tissue Chips display, scoring and data entry options are shown. UL: A reconstructed colon cancer microarray is shown with thumbnails of core images displayed in a rectangular array matching the layout on the slide. A mouse click on one of the thumbnails pops up the full screen version of the core image (ML). Any area of interest can be viewed at native magnification (LL). From both full screen (ML) and native magnification (LL) views, the properties of the core can be annotated using the properties button. UR: Automated core segmentation performed on images acquired on the commercially available system Aperio system. Top: Reduced-resolution (100:1) montage of all fields of view. Bottom: Screen shot of Tissue Chips, where the cores have been automatically extracted from the montage and properly arrayed for future analysis.

10×0.5 NA objective capable of $d = 0.6\lambda/NA = 0.6$ μm optical resolution at wavelength $\lambda = 0.5$ μm) was collected for each 600 μm diameter core after the cores were located by a scan with 7 μm sampling [29]. If the technique allows only for one image of each core, TMAs with larger cores would be either more incompletely imaged under the same conditions, or more undersampled. Acquiring one image per core may limit flexibility for acquiring high resolution images of each entire core. On the other hand, a low resolution scan may not require autofocus if the only goal is to grossly locate the tissue sections. Low resolution scans can miss small fragments of tissue and relocating for only a single image might miss some large fragments of broken cores in TMAs. For some, it may be primarily a matter of convenience to have images representing all of the tissue section areas on the slide. However, if this convenience can be achieved at a reasonable speed, then there is an advantage to scanning the entire area once at final resolution because all raw images necessary to completely reconstruct fragmented sections are acquired. Scanning the entire area is also more general because it allowed creation of an algorithm that makes no assumptions about the pattern of the tissue sections. The software we developed will group and reconstruct a single image from many images of tissue sections of any size and positioning.

With automated image collection now a reality, the next step is to automate scoring. For IHC, automated scoring is performed by densitometry of the expression stain, which must first be separated from the hematoxylin signal. We recently demonstrated the utility of independent components analysis (ICA) and nonnegative matrix factorization (NMF) for unsupervised multispectral decomposition of DAB and hematoxylin [30]. Unsupervised decomposition, or unmixing, creates pure DAB and hematoxylin images from which DAB densitometry can be performed directly. Since these techniques do not require spectral calibration of the different stains, they offer a powerful advantage for automated densitometry.

With densitometry automated, the final step is to differentiate normal glandular and stromal elements from cancer tissue to obtain independent expression scores. Camp *et al.* used polyclonal anti-cytokeratin and anti-α-catenin immunofluorescence to automatically differentiate breast and colon cancer cells from stroma and obtain colocalized integrated intensity scores specific to cancer [29]. It was assumed that cancer biopsies contained only cancer and stromal tissue, without contaminating normal tissue and stroma (*personal communication*). However, the specificity of cancer versus normal glandular staining has not been explored systematically on a cell-by-cell basis. While cytokeratins are widely used to detect breast cancer cells in tumor tissue, lymph nodes, bone marrow, and blood, it has been known for many years that this staining technique is not 100% specific. Primary adenocarcinomas and their metastases have demonstrated greater than 95% tumor cell staining [32]. However, in this same study, benign breast lesions showed a "mosaic pattern" of staining. In another study of surface antigen expression of breast cancer cells in micrometastases in bone marrow, the surface antigens were shown to have different expression patterns than the primary tumor [33]. Therefore, while it is possible that a specific immunostain, or set of immunostains will be highly correlated with tumor tissue in a given tissue section or TMA core, errors are likely due to the inherent heterogeneity of cancer, and this error rate is as yet unknown. Given the frequent admixture of normal epithelium and invasive cancer cells, it is possible that the subset of tumor tissue in a particular TMA core will exhibit little or no staining even if the bulk of the tumor is marker positive. Nevertheless, this approach enables complete automation with compelling simplicity and appropriate cocktails of antibodies may become sophisticated enough to differentiate tumor, normal tissue, and stroma in TMA applications, especially given the powerful statistical advantage of automating the reading of large TMA cohorts. The Camp *et al.* study, for example, showed that automated analysis had better reproducibility ($R = 0.824$ versus $R = 0.732$) and slightly outperformed manual analyses at predicting patient survival based on estrogen receptor in breast cancers. With the addition of unsupervised spectral unmixing [30], it should be possible to apply the same principles to immunohistochemistry stains.

The predictive value of the data analysis and the speed of the system will dominate the performance criteria. In addition, the extent of automation is an important variable. Manual interventions generate substantial speed penalties. As automation improves, the speed should be reported as the time required for analyzing an appropriate test set of TMA slides (say 20 slides of 50–2500 core arrays) to ensure inclusion of all required steps. Fully automated techniques such as those used by Camp *et al.* [29] will likely provide the fastest, most consistent performance. The combination of comprehensive scanning and Tissue Chips image reconstruction/display reported here contributes the complete front-end automation required for enabling scanning of a large number of slides from a slide loader and loading completely reconstructed core images into a database. The resulting images can be read manually and enable research on the development of methods for fully automated scoring, including texture segmentation methods that may separate normal glandular elements, cancer and stromal tissue in classically stained sections. The flexibility of scanning fluorescence (and reflection) or brightfield images enables exploration of other staining and analyses techniques. Given that pathologists prefer brightfield stains for diagnoses, this flexibility may be important for fast comparisons between manual (brightfield) and automated scoring techniques even if based on fluorescence.

An important practical concern for scanning the area exhaustively is speed. The instrument used here [8] can achieve a peak speed of about 3 fields/s. With a scan area of 5,000 fields, it would acquire the images and reassemble them into the database in about 30 min with an RGB camera. Aperio's system operates in a continuous-scanning fashion that decreases scanning time to a few minutes, but requires setup time for each slide to prescan a grid of focus points to guide continuous scanning. The number of grid points needed and the prescan time required depends on the depth of field (or NA) of the objective. We have demonstrated continuous scanning with closed feedback autofocus [19], [20] that eliminates the need for prefocussing and will also reduce the scanning time to a few minutes. The existence of

these approaches for scanning TMA slides at high speed bode well for the convenience and flexibility of exhaustively scanning the whole area and allowing the software to locate tissue sections arrayed in any pattern, whether regular or irregular.

V. CONCLUSION

Databases combining manually acquired TMA images and study results with clinical diagnostic information have been reported [12], [31]. Our automated software methods will enable routine processing of automatically reconstructed super-bitmaps, or montaged images, of each TMA core or tissue section by the computer, within a database, for permanent archiving. For pathologists, this can be linked to a reporting mechanism to automate routine diagnostic activities. Persistence in a clinical database will enable real-time recall of images and reports with subsequent tissue samples for diagnostic comparisons of tumor progression. Linking the images, experimental methods (stains and molecular targets) and scores to the clinical information will greatly simplify TMA research.

Advances in automation are important for reading large numbers of tissue samples. Automation will facilitate creating databases of information easily accessible by researchers. We have presented a set of automated preprocessing, reconstruction, display and organization (archiving for database storage) techniques that automate tissue section and TMA scanning for visualization at high resolution with fluorescence or brightfield microscopy. Combined with additional image processing methods, including unsupervised multispectral stain-component decomposition for automated densitometry [30], these techniques bring us closer to the level of automation required for unattended scanning of large numbers of slides at the full range of magnifications available on light microscopes.

REFERENCES

[1] J. Kononen, L. Bubendorf, and A. Kallioniemi, "Tissue microarrays for high-throughput molecular profiling of tumor specimens," *Nature Medicine*, vol. 4, pp. 844–847, 1998.

[2] M. S. Skacel, B. Skilton, J. D. Pettay, and R. R. Tubbs, "Tissue microarrays: A powerful tool for high-throughput analysis of clinical specimens," *Applied Immunohistochemistry and Molecular Morphology*, pp. 1–6, 2002.

[3] R. L. Camp, C. A. Lori, and D. L. Rimm, "Validation of tissue microarray technology in breast carcinoma," *Laboratory Investigation*, vol. 80, no. 12, pp. 1943–1952, 2000.

[4] D. Zhang, M. Salto-Tellez, T. C. Putti, E. Do, and E. S. Koay, "Reliability of tissue microarrays in detecting protein expression and gene amplification in breast cancer," *Modern Pathology*, vol. 16, no. 1, pp. 79–84, Jan. 2003.

[5] M. A. Rubin, R. Dunn, M. Strawderman, and K. J. Pienta, "Tissue microarray sampling strategy for prostate cancer biomarker analysis," *Amer. J. Surgical Pathology*, vol. 26, no. 3, pp. 312–319, Mar. 2002.

[6] J. H. Price, "Scanning Cytometry for Cell Monolayers," Ph.D. dissertation, Dept. of Bioengineering, Univ. of California, San Diego, 1990.

[7] J. H. Price, E. A. Hunter, and D. A. Gough, "Accuracy of least squares designed spatial FIR filters for segmentation of images of fluorescence stained cell nuclei," *Cytometry*, vol. 25, pp. 303–316, 1996.

[8] M. Bravo-Zanoguera, B. Massenbach, A. L. Kellner, and J. H. Price, "High-performance autofocus circuit for biological microscopy," *Review of Scientific Instruments*, vol. 69, no. 11, pp. 3966–3977, 1998.

[9] S. Heynen, D. A. Gough, and J. H. Price, "Optically stabilized mercury vapor short arc lamp as uv-light source for microscopy," in *SPIE-Int. Soc. Opt. Eng. Proc. Optical Diagnostics of Biological Fluids and Advanced Techniques Analytical Cytology*, vol. 2982, 1997, pp. 430–434.

[10] S. Bajaj, J. B. Welsh, R. C. Leif, and J. H. Price, "Ultra-rare-event detection performance of a custom scanning cytometer on a model preparation of fetal nRBCs," *Cytometry*, vol. 39, pp. 285–294, 2000.

[11] S. Heynen and J. H. Price, "Evaluation of scanning cytometer fluorometry performance," in *SPIE-Int. Soc. Opt. Eng. Proc. Optical Diagnostics of Biological Fluids and Advanced Techniques Analytical Cytology*, vol. 3604, San Jose, CA, 1999, pp. 237–242.

[12] S. Manley, N. R. Mucci, A. M. De Marzo, and M. A. Rubin, "Relational database structure to manage high-density tissue microarray data and images for pathology studies focusing on clinical outcome," *Amer. J. Pathology*, vol. 159, no. 3, pp. 837–843, 2001.

[13] M. Krajewska, J. M. Zapata, I. Meinhold-Heerlein, H. Hedayat, A. Monks, H. Bettendorf, A. Shabaik, L. Bubendorf, O. P. Kallioniemi, H. Kim, G. Reifenberger, J. C. Reed, and S. Krajewski, "Expression of Bcl-2 family member bid in normal and malignant tissues," *Neoplasia*, vol. 4, no. 2, pp. 129–140, 2002.

[14] J. Crookham and R. Dapson, *Hazardous Chemicals in the Histopathology Laboratory*, 2nd ed., Springfield, OH: Anatech, 1991.

[15] D. Sheehan and B. Hrapchak, *Theory and Practice of Histotechnology*, 2nd ed., Columbus, OH: Battelle, 1980, pp. 42, 46–43, 47.

[16] M. Rentrop, B. Knapp, H. Winter, and J. Schweizer, "Aminoalkylsilane—Treated glass slides as support for *in situ* hybridization of keratin cDNAs to frozen tissue sections under varying fixation and pretreatment conditions," *J. of Histochemistry*, vol. 18, pp. 271–276, 1986.

[17] S. Krajewski, M. Krajewska, L. M. Ellerby, K. Welsh, Z. Xie, Q. L. Deveraux, G. S. Salvesen, D. E. Bredesen, R. E. Rosenthal, G. Fiskum, and J. C. Reed, "Release of caspase-9 from mitochondria during neuronal apoptosis and cerebral ischemia," in *Proc. Nat. Acad. Sci.*, vol. 96, 1999, pp. 5752–5757.

[18] S. Krajewski, M. Krajewska, A. Shabaik, T. Miyashita, H. G. Wang, and J. C. Reed, "Immunohistochemical determination of *in vivo* distribution of Bax, a dominant inhibitor of Bcl—2," *Amer. J. Pathology*, vol. 145, pp. 1323–1336, 1994.

[19] M. Bravo-Zanoguera, B. Massenbach, and J. H. Price, "Automatic on-the-fly focusing for continuous image acquisition in high-resolution microscopy," in *SPIE-Int. Soc. Opt. Eng. Proc. Optical Diagnostics of Biological Fluids and Advanced Techniques Analytical Cytology*, vol. 3604, San Jose, 1999, pp. 243–252.

[20] L. K. Nguyen, M. Bravo-Zanoguera, A. L. Kellner, and J. H. Price, "Magnification corrected optical image splitting module for simultaneous multiplanar acquisition," in *Proc. SPIE-Int. Soc. Opt. Eng., Optical Diagnostics of Living Cells III*, vol. 3921, San Josa, CA, 2000, pp. 31–40.

[21] A. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least squares procedures," *Analytical Chemistry*, vol. 36, pp. 1627–1639, 1964.

[22] N. Ramesh, J. H. Yoo, and I. K. Sethi, "Thresholding based on histogram approximation," in *IEEE Proc., Vision, Image Signal Process.*, vol. 142, 1995, pp. 271–279.

[23] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. 9, pp. 62–66, 1979.

[24] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification,* 2nd ed. New York: Wiley, 2001.

[25] R. C. Gonzalez and P. Wintz, *Digital Image Processing,* 2nd ed. Reading, MA: Addison-Wesley, 1987.

[26] D. G. Luenberger, *Linear and Nonlinear Programming,* Reading, MA: Addison-Wesley, 1984.

[27] K. Alsabti, S. Ranka, and V. Singh, "An efficient k-means clustering algorithm," presented at the IPPS/SPDP Workshop on High Performance Data Mining, Orlando, Florida, 1998.

[28] S. Inoue and K. Spring, *Video Microscopy—The Fundamentals*, 2nd ed., New York: Plenum Press, 1997.

[29] R. L. Camp, G. G. Chung, and D. L. Rimm, "Automated subcellular localization and quantification of protein expression in tissue microarrays," *Nature Medicine*, vol. 8, no. 11, pp. 1323–1327, 2002.

[30] A. Rabinovich, S. Agarwal, C. Laris, J. H. Price, and S. Belongie, "Unsupervised color decomposition of histologically stained tissue samples," in *Advances in Neural Information Processing Systems*, S. Thrun, L. Saul, and B. Schölkopf, Eds. Cambridge, MA: MIT Press, 2004, vol. 16.

[31] C. L. Liu, W. Prapong, Y. Natkunam, A. Alizadeh, K. Montgomery, C. B. Gilks, and M. van de Rijn, "Software tools for high-throughput analysis and archiving of immunohistochemistry staining data obtained with tissue microarrays," *Amer. J. Pathology*, vol. 161, pp. 1557–1565, 2002.

[32] J. Bártek, J. Bártková, J. Schneider, J. Taylor-Papadimitriou, J. Kovařík, and A. Rejthar, "Expression of monoclonal antibody-defined epitopes of keratin 19 in human tumours and cultured cells," *Eu. J. Cancer and Clinical Oncology*, vol. 22, no. 12, pp. 1441–52, 1986.

[33] S. Braun, F. Hepp, H. L. Sommer, and K. Pantel, "Tumor-antigen heterogeneity of disseminated breast cancer cells: Implications for immunotherapy of minimal residual disease," *Int. J. Cancer*, vol. 84, no. 1, pp. 1–5, 1999.

[34] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements Of Statistical Learning,* 1st ed.   New York: Springer-Verlag, 2001.

Ahmed Shabaik received the M.D. degree from Cario University, Egypt, in 1977.

He is Professor of Pathology at the University of California at San Diego, and is certified by the American Board of Pathology in AP, CP, and Cytopathology. His areas of expertise are male genitourinary pathology and fine needle aspiration biopsy performance and interpretation. He has experience in cell image analysis, particularly with quantitation of DNA and ploidy analysis. In addition to his clinical duties, he teaches pathology residents, urology residents, and medical students.

Dr. Shabaik is a member of many professional organizations including the American Society of Clinical Pathologists (ASCP), United States and Canadian Academy of Pathology (USCAP), International Society of Urologic Pathology (ISUP), and the American Society of Cytopathology (ASC). He is the recipient of the Best Anatomic Pathology Teacher award for 2002.

Andrew Rabinovich received the B.S. degree in computer science and engineering (CSE) from the University of California at San Diego in 2003, where he received two Chancellor's Research Fellowships.

Currently, he is a Ph.D. student at the University of California at San Diego in the CSE department. His research interests include computer vision, perceptual organization, and automated cancer detection.

Stan Krajewski received the M.D. degree at the Gdansk Medical Academy, Poland, and the Ph.D. degree from the Institute of Medical Research, Polish Academy of Sciences, Warsaw, Poland in 1979. He received his postdoctoral training in clinical and experimental immunopathology at the Department of Immunopathology, National Institute of Health, Warsaw, and in clinical and experimental neuropathology at the Department of Neuropathology, Institute of Medical Research and Nencki Institute, Warsaw.

He earned the title of Professor of Medical Science (Neurology and Neuropathology) and was appointed as the Head of the Laboratory of Neuroimmunopathology, Department of Neuropathology, Institute of Medical Research, Polish Academy of Sciences, Warsaw, an appointment he maintains to this day. Since 1992, he has been working in the United States and is currently a Research Associate Professor at The Burnham Institute in the Program on Apoptosis, Cell Death and Aging Research. His research interests combine the profiles of immunohistological investigation of novel genes and proteins in tissues with new, high throughput technologies like tissue microarrays and high throughput microscopy. He has over 25 years of experience in diagnostic histopathology and immunohistochemistry, and has published over 200 scientific papers and participated in eight book chapters.

Stephen M. Hewitt received the B.A. degree in philosophy from The Johns Hopkins University, the M.D. degree from the University of Texas Medical School, Houston, and the Ph.D. degree in genetics from The University of Texas Graduate School of Biomedical Sciences, Houston.

He did his residency in anatomic pathology at the National Cancer Institute (NCI) and is Board Certified in Anatomic Pathology. In 2000, he initiated the Tissue Array Research Program at the NCI, of which he is currently Chief. He has published articles that cover a diversity of fields, with emphasis on angiogenesis, tissue microarrays, and the use of biomarkers in clinical medicine.

Maryla Krajewska received the M.D. degree from the Gdansk Medical Academy, Poland, in 1981. She received her postdoctoral training in neurosurgery and neuropathology at University of Duesseldorf, Germany, in 1982–1985.

From 1986 to 1992, she was a Research Associate at the Psychoneurological Institute of Polish Ministry of Health, Warsaw, Poland. In 1996, she became associated with The Burnham Institute where she is currently a Senior Staff Scientist. She has over 70 peer reviewed publications.

Serge Belongie (S'93–M'01) was born in Sacramento, CA, in 1974. He received the B.S. degree (with honor) in electrical engineering from the California Institute of Technology, Pasadena in 1995, and the M.S. and Ph.D. degrees in electrical engineering and computer sciences (EECS) from the University of California at Berkeley in 1997 and 2000, respectively. While at Berkeley, his research was supported by a National Science Foundation Graduate Research Fellowship.

He is a Co-Founder of Digital Persona, Inc., and the principal architect of the Digital Persona fingerprint recognition algorithm. He is currently an Assistant Professor in the Computer Science and Engineering Department at the universtiy of California at San Diego. His research interests include computer vision and pattern recognition.

Dr. Belongie is a recipient of the NSF CAREER Award and the Alfred P. Sloan Research Fellowship. In 2004, *MIT Technology Review* named him to the list of the 100 top young technology innovators in the world (TR100).

John C. Reed received the M.D. and Ph.D. degrees from the University of Pennsylvania School of Medicine, in 1986, when he also received his clinical and postdoctoral training.

In 1989, he was appointed Assistant Professor, Department of Pathology and Laboratory Medicine at University of Pennsylvania School of Medicine, and Assistant Director, Laboratory of Molecular Diagnosis at the Hospital of the University of Pennsylvania. He was recruited to the Burnham Institute for Medical Research, in 1992. He served as the Institute's Scientific Director from 1995 to 2001, and was appointed President and CEO on January 1, 2002. He is the author of over 600 papers and has published more articles on apoptosis and cell death in the past decade than any other scientist world-wide, according to the Institute for Scientific Information. During 1996–1998, he ranked first among biomedical researchers world-wide for the impact of his research. He is the recipient of numerous awards and honors, and serves on the editorial boards of over ten scientific journals. He is Scientific Cofounder of IDUN Pharmaceuticals, Inc., which was recently acquired by Pfizer, and is Inventor of the anticancer drug GenasenseTM.

Jeffrey H. Price (A'05) received the M.D. degree from Loma Linda University, Loma Linda, CA, in 1985, and the Ph.D. degree in bioengineering from the University of California, San Diego, in 1990, where he also received his postdoctoral training.

In 1993, he was appointed Assistant Project Scientist and, in 1996, Associate Research Scientist. From 1994 to 2004 he directed the NSF-Whitaker Quantitative Imaging and Confocal Microscopy Resource in Bioengineering at UCSD. In 1999 he founded Q3DM Inc, which marketed his research group's high throughput microscopy inventions for cell-image-based screening which was purchased by Beckman Coulter, Inc. in 2003. In 2004 he Co-Founded Vala Sciences Inc, which develops and markets software and kits for cell-image-based assays. He was recruited to the position of Associate Professor at The Burnham Institute in 2004, where he is a member of the Signal Transduction and Stem Cells and Regeneration Research Programs. With its collaborators, his laboratory is focused on automated analytical microscopy research for tissue proteomics, tracking of cell migration and differentiation, and rare cell diagnostics.