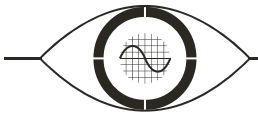


Lehrstuhl für Bildverarbeitung
Institute of Imaging & Computer Vision

Fast Virtual Destaining of Immunocytological Specimens

André Bell and David Friedrich and Kraisorn Chaisaowong and Till Braunschweig and Ruth Knüchel-Clarke and Til Aach

Institute of Imaging and Computer Vision
RWTH Aachen University, 52056 Aachen, Germany
tel: +49 241 80 27860, fax: +49 241 80 22200
web: www.lfb.rwth-aachen.de

source undefined

BibTeX:

```
@inproceedings{BEL11a,
title = {Fast Virtual Destaining of Immunocytological Specimens},
publisher = {IEEE},
year = {2011},
month = {March 30 -- April 2},
address = {Chicago},
note = {914--917},
author = {Andr\'{e} Bell and David Friedrich and Kraisorn Chaisaowong and Till Braunschweig and Ruth Kn{\"u}chel-Cl
```

FAST VIRTUAL DESTAINING OF IMMUNOCYTOLOGICAL SPECIMENS

A. A. Bell[⋆]*, D. Friedrich*[⋆]*, K. Chaisaowong*[⋆]*, T. Braunschweig*[†]*, R. Knüchel-Clarke*[†]*, T. Aach*[⋆]

[⋆] Institute of Imaging and Computer Vision, RWTH Aachen University, D-52074 Aachen
[†] Institute of Pathology, University Hospital Aachen, D-52074 Aachen

ABSTRACT

Microscopy-based diagnosis of certain diseases or infections, e.g. with human papilloma viruses (HPV) for the identification of high risk patients for cervical cancer, relies more and more often on immunocytochemical marker stains. These markers stain cells that exhibit a particular protein. In addition to one or more marker stains, pathologists need to simultaneously assess the morphology of the cell. Therefore the specimens are commonly also stained with a counter stain.

For a reliable computer assisted analysis of such multiple-stained images a virtual destaining of all but one of the stains is required. This can be achieved by a specific color transformation. The transformation depends on the staining colors, and is therefore in general non-orthogonal.

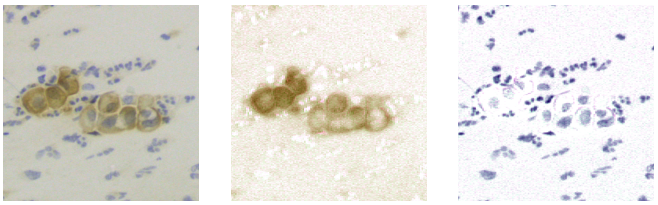
In this paper we analyze the performance of this color space conversion. Based on a reformulation of the problem we develop faster implementations. We analyze the complexity of these implementations and demonstrate that with a marginal additional memory footprint the virtual destaining can be calculated $5 - 10$ times faster than the conventional method. This, on the one hand, speeds up the automated analysis. More importantly, it enables pathologists to view the virtually destained slides and change the color transformation interactively at high frame rate.

***Index Terms*—** immunocytochemical marker, cytology, color separation, virtual destaining, cancer diagnosis

1. INTRODUCTION

For certain types of cancer and their precursor stages it could be shown that these lesions can be highlighted on the cellular level by disease-specific proteins within the cells. To this end, cell specimens are taken from the patient, stained and investigated with a microscope. The visualization of such proteins depends on highly protein-sensitive and protein-specific stains. Over the recent years an increasing number of such so-called immunocytochemical marker stains have been developed.

The project was supported by the excellence initiative of the German federal and state governments. We thank Prof. R. H. Jansen, german director of the Thai-German Graduate School of Engineering, KMUTNB, for fruitful discussion and cooperation.



(a) (b) (c)

Fig. 1. Example image obtained from a cervical brush smear, a so-called PAP smear. The slide has been stained with a counter stain (hematoxylin, blue) and the p16-antibody (brown) (a). After a stain specific color separation two individual images demonstrating only the immunocytochemical marker stain (b) or the counter stain (c) are obtained.

For example, nearly all cervical cancers are caused by human papilloma viruses (HPV) [1]. This infection in turn is directly linked to an upregulation of a protein. A specific corresponding antibody is docked to this protein within the infected and altered cell and stained with immunocytological techniques. For the case of cervical brush smears, so-called PAP smears, it has been shown that the p16-antibody identifies infections by HPV with high sensitivity [2]. Since the cell specimens are obtained non-invasively with a brush, the immunocytochemical marker enables reliable identification of high-risk patients in screening scenarios. Only patients with HPV infection have to be examined further by cytopathological means and resampling in short time frame. Still, to cope with the high load of specimen, an automation for the investigation of misroscopical slides is of great benefit.

Generally the cells are stained by a so-called counter stain and one or more immunocytochemical marker stains. The former allows to investigate the cell morphology. Moreover, the counter stain visualizes all cells in the specimen. Since only a fraction – if at all – of cells are marker positive, the counter stain enables a reliable autofocus even in case of marker negative fields of view (FOV) [3, 4]. For immunocytochemical marker stains with a color expression in the cytoplasm, the counter stain allows to identify the nucleus within the cytoplasm, enabling to distinguish stained cells from artifacts with localized unspecific color precipitation without a nucleus.

Automated detection of marker positive cells, however, is facilitated if marker-positivity and nuclei morphology could

be individually investigated. A new form of investigation would be available, if the pathologist could also investigate virtually destained slides. In [5], a method is proposed which algorithmically separates images obtained from microscopical slides into stain-specific channels (see Fig. 1). The stains are separated by estimating their optical densities and decomposing the image into stain specific contributions. The contribution from undesired stains are removed and a new image is reconstructed, now only showing the desired stain.

Processing large images in this manner is time consuming. In this paper we discuss performance improvements for the algorithm proposed by Ruifrok [5]. Using a lookup table (LUT) for the computation of the optical densities increases the performance by a factor of two. Yet the whole color separation process can be precomputed and stored in a LUT, which increases the performance by a factor of 10. However, the calculation of the latter LUT is expensive in terms of time and memory usage, and it has to be recomputed after changes of the lighting conditions and changes in the stain colors. We reformulate the algorithm and show that high speedup and fast update of a LUT can be achieved *simultaneously*. Performance and scalability of the algorithms are analyzed and measurement results about the performance are given.

2. METHOD

The image formation process of specimens in bright field light microscopy can be described by the Lambert-Beer law for the attenuation of light passing through the specimen. Let the sensor response to directly incident white light, not attenuated by any cell, given as I_w. This light passes through the material of the specimen. Depending on the optical density of the material and the stain the light attenuation is described by

$$f_I(OD) := I_w \cdot e^{-OD}. \tag{1}$$

The optical density OD in turn is linearly related to the concentration of the stain in the specimen. Therefore, Lambert-Beer's law (1) allows to measure the *overall* optical density of particular stained characteristics of the specimens.

2.1. Virtual destaining of immunocytological specimen

In the case of multiple stained specimens it is required to separate the images into stain specific channels. This enables measuring the optical density of each stain in the specimen *individually*. To this end Ruifrok [5] suggested a *non-orthogonal* color space transformation based on the measured color responses of the individual stains. For every acquired $I = (I_R, I_G, I_B)^T$ vector, the corresponding optical densities OD are computed by inverting Eq. (1) given by

$$f_{OD}(I) := f_I(OD)^{-1} = -\log\left(I./I_w\right), \tag{2}$$

where $./$ is the component-wise division. Using this equation, RGB values corresponding to the individual stains can be transformed to their optical densities. The optical density vectors u, d, and n define the stain matrix $\mathbf{M}_s$ as

$$\mathbf{M}_s := (d, u, n). \tag{3}$$

If only two stains are present on a slide, e.g., an immunocytochemical marker d (for desired) and counter stain u (undesired), the two vectors have to be complemented by a third vector n to span the RGB color space. For that case, we choose $n = u \times d$ as the third spectral component. The matrix $\mathbf{M}_s$ maps concentrations to optical densities. Consequently, the inverse $\mathbf{M}_s^{-1}$ of $\mathbf{M}_s$ transforms optical densities to the individual stain concentrations c_d, c_u, c_n:

$$c = \begin{pmatrix} c_d \\ c_u \\ c_n \end{pmatrix} = \mathbf{M}_s^{-1} f_{OD}(I). \tag{4}$$

To virtually destain the image from one of the components (e.g., u) we set the corresponding concentration $c_u = 0$ by multiplication with the destain matrix $\mathbf{M}_{dn}$ given by

$$\mathbf{M}_{dn} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{5}$$

The destain matrices $\mathbf{M}_{un}$ and $\mathbf{M}_{du}$ are defined accordingly. A virtually destained slide can hence be calculated from the new concentration vector by reprojecting into optical densities with $\mathbf{M}_s$ and converting from OD values to RGB values. The full virtual destaining is described by

$$\hat{I}_{dn} = f_I\left(\mathbf{M}_s \mathbf{M}_{dn} \mathbf{M}_s^{-1} f_{OD}(I)\right). \tag{6}$$

This can be written more compactly as

$$\hat{I}_{dn} = f_I\left(\mathbf{M} f_{OD}(I)\right), \tag{7}$$

where $\mathbf{M} = \mathbf{M}_s \mathbf{M}_{dn} \mathbf{M}_s^{-1} = m_{r,c}$ is a 3×3 matrix for RGB images and $r, c \in [0, 1, 2]$ are the row and column indices of the entries $m_{r,c}$ in $\mathbf{M}$.

This virtual destaining technique is well suited for stains with OD vectors d and u which are linearly independent and non-zero. Otherwise, and also when d and u are close to linear dependency, the inverse of the matrix $\mathbf{M}_s$ is not defined or badly conditioned.

2.2. Fast implementations

In the original implementation acc. to Eq. (7), which we denote as "*Ruifrok destaining*" (A), all image pixels are individually converted to their optical densities, the undesired stain is removed and the result is transformed back to RGB values.

To speed up this implementation we first note that the function $f_{OD}(I)$ maps integer values to floating point numbers enabling an implementation with a lookup table (LUT). Given the sensor response I_w to the incident white light, all

Table 1. The table gives the timing results for virtual destaining and calculation of the required LUTs on 5 Megapixel images with 8 bit per channel for all different implementations. Moreover the complexity of the LUT computation is provided.

		double (64 bit)	complexity
(A)	Ruifrok destaining	2660 ms	
(B)	OD-LUT calculation	< 1 ms	$\mathcal{O}(N_c \cdot 2^{N_b})$
	Ruifrok destaining OD-LUT	1460 ms	
(C)	Full-LUT calculation	2920 ms	$\mathcal{O}(N_c \cdot 2^{N_b \cdot N_c})$
	Ruifrok destaining Full-LUT	253 ms	
(D)	Power-LUT calculation	< 1 ms	$\mathcal{O}(N_c^2 \cdot 2^{N_b})$
	Power destaining	545 ms	

optical densities which the sensor is able to capture can be pre-calculated with $f_{OD}(I)$ from Eq. (2). For an 8 bit image, for instance, these are 256 floating point numbers for every color channel. The optical densities from the LUT are then used for the further computation of the virtually destained image in Eq. (7). We call this implementation "*Ruifrok destaining OD-LUT*" (B). Note, however, that for transforming the optical densities back to RGB values with $f_I(OD)$, a straightforward LUT implementation cannot be given, since this function maps floating point numbers to floating point numbers.

Still, since Eq. (7) maps each RGB vector to a floating point RGB vector, it is possible to pre-calculate the destaining results for *all* RGB vectors. The calculation of this LUT depends on I_w, u, d, and n. This LUT implementation maps the input for Eq. (7) directly to the color image output and consequently requires the computation of Eq. (7) for all possible combinations of camera sensor responses. We denote this implementation as "*Ruifrok destaining Full-LUT*" (C).

The computation of the Full-LUT has a high complexity (see Section 3). To achieve a lower complexity we reformulate the color separation algorithm. First, we rewrite Eq. (7) with f_{OD} and f_I to

$$\hat{I}_{dn} = I_w \cdot e^{\mathbf{M} \log(I./I_w)}. \tag{8}$$

We now write Eq. (8) for one channel r giving

$$\hat{I}_{dn,r} = I_{w,r} \cdot e^{\sum_{c=0}^{2} m_{r,c} \log(I_c/I_{w,c})}. \tag{9}$$

This can be written as

$$\hat{I}_{dn,r} = I_{w,r} \cdot \prod_{c=0}^{2} e^{\log(I_c/I_{w,c}) m_{r,c}}. \tag{10}$$

As the logarithm is the inverse of the exponential, the equation can now be simplified to

$$\hat{I}_{dn,r} = I_{w,r} \prod_{c=0}^{2} \left(\frac{I_c}{I_{w,c}} \right)^{m_{r,c}}. \tag{11}$$

The exponents $m_{r,c}$ in Eq. (11) are the elements of the matrix $\mathbf{M}$ and therefore depend on the optical densities d, u and n of the stains. The values for every product term $(I_c/I_{w,c})^{m_{r,c}}$ can now be pre-calculated and stored in a LUT. For a RGB image, for instance, this requires the computation of 9 LUTs. The RGB values for a virtually destained slide are then computed by looking up the factors in Eq. (11) in the corresponding LUT and multiplying them with each other and I_w. We denote this algorithm as "*Power destaining*" (D).

3. ANALYSIS OF THE IMPLEMENTATIONS

Let N be the number of pixels in the image, N_c the number of color channels, and N_b the number of bits of the quantized camera output (yielding 2^{N_b} possible sensor outputs).

The overall algorithm as given by Eq. (7) is $\mathcal{O}(N)$. Therefore, all implementations (A)-(D) given in Section 2 are implementational improvements (measured in Section 4) and do not change the order of the algorithm.

However, the computation of the lookup tables differs for the presented approaches and thus leads to different computational complexities and memory requirements for the LUTs.

(A) *Ruifrok destaining*: This implementation does not exploit LUTs and thus requires no additional computation time or memory. However, this implementation has the computationally most expensive inner loop.

(B) *Ruifrok destaining OD-LUT*: For every of the 2^{N_b} possible camera outputs, an optical density has to be computed. As the I_w differs for every color channel, such a LUT has to be calculated for every color channel, thus the complexity is $\mathcal{O}(N_c \cdot 2^{N_b})$.

(C) *Ruifrok destaining Full-LUT*: The number of all possible sensor combinations, for which the LUT is precomputed, is given by $\left(2^{N_b}\right)^{N_c}$. Every entry of the LUT contains N_c values, therefore the complexity and memory footprint for the Full-LUT is $\mathcal{O}(N_c \cdot 2^{N_b \cdot N_c})$. For instance for an 8 bit RGB camera, the Full-LUT consists of $256^3 \approx 16.7M$ entries of floating point RGB triplets. Consequently, this Full-LUT using double precision floating point (8 Byte) requires $256^3 \cdot 3 \cdot 8$ Byte $\approx$ 400 MB of memory.

(D) *Power destaining*: For every exponent in Eq. (11) (or equivalently, for every entry of the $N_c \times N_c$ matrix $\mathbf{M}$), a LUT has to set up, which in turn consists of 2^{N_b} entries. Thus the complexities is $\mathcal{O}(N_c^2 \cdot 2^{N_b})$.

4. EXPERIMENTS AND RESULTS

We acquired images of immunocytochemically stained cervical smears, which have been stained with hematoxylin and the p16-marker (REAL-Kit from DAKO, Glostrup, Denmark with DAB (diaminobenzidine) as brown colorization), with a Nikon Ti-E microscope, $10x$ objective (NA $= 0.3$) and equipped with a Baumer TXG50c 5 Megapixel RGB camera. All images were acquired in 8 bit mode and all algorithmic variants have been implemented in C with double precision (64 bit) floating point. The calculations have been carried out on a notebook with Intel Core2 Duo CPU (T9400) at 2.53GHz and 4GB RAM. Timing results are given as the mean calculation time over 100 runs and are given in Table 1. Image results are shown in Fig. 1 and Fig. 2.

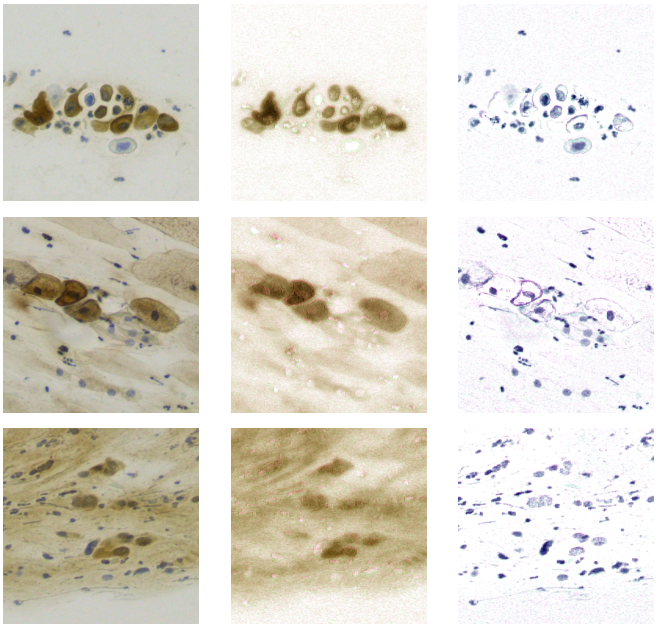
5. DISCUSSION

In this paper we have analyzed the color separation algorithm by Ruifrok [5]. We reformulated the algorithm and demonstrated that for the microscope with imaging described by Lambert-Beer's law, virtual destaining can be achieved by calculating a product of one factor for each color channel.

The fastest color separation can be achieved with LUT for every possible RGB color vector. However, this Full-LUT requires a huge amount of memory even for common 3-channel 8 bit images. Since the memory footprint increases exponentially with the channel count and the bit depth, it becomes infeasible for multispectral cameras or higher bit depth. The *power destaining*, on the other hand, only requires a moderate amount of memory for the lookup tables and achieves almost the same speedup. Moreover, recalibrating the color separation or when interactively adapting the color vectors of the underlying stains is fastest with the *power destaining* implementation, since for this setting the overall time of LUT calculation and color separation has to be as low as possible.

The joint speedup of the *power destaining* implementation for calibration and color separation thus enable *interactive* color selection and live screen display of virtually destained slides. To this end, the pathologist selects stain colors and white point by clicking on the corresponding points in a live image from the photo microscopic system. The virtually destained slide with the recalibrated parameters is then shown with high frame rate. For instance, we achieve a frame rate of $\approx 7 fps$ at screen resolution (1.2 Megapixel).

All algorithms are suitable for GPU implementation, which we will investigate next. Finally, we point out that the current algorithm is given for the special case of transmission light microscopy imaging under Lambert-Beer's law.



(a) (b) (c)

Fig. 2. This figure shows more examples of the virtual destaining. Column (a) are the images acquired by the camera. Counter stain and marker stain are mixed. In columns (b) and (c) the marker stain and counter stain after virtual destaining are shown.

However, for the case of reflective imaging an adapted algorithm has been proposed [6]. Our speed improvements can be straightforwardly applied to that algorithm as well. In this case, the product in Eq. (11) just becomes a sum.

6. REFERENCES

[1] M. Schiffman and P.E. Castle, "The promise of global cervical cancer prevention," *N Engl J Med*, vol. 353, 2101–2104, 2005.

[2] K. J. Denton, C. Bergeron, and et al., "The sensitivity and specificity of p16(ink4a) cytology vs HPV testing for detecting high-grade cervical disease in the triage of ASC-US and LSIL pap cytology results.," *Am J Clin Pathol*, vol. 134(1), 12–21, 2010.

[3] A. A. Bell, T. E. Schneider, and et al., "Fully automatic screening of immunocytochemically stained cytopathological specimens for early cancer detection," in *SPIE Medical Imaging*, 2007, vol. 6514, 651431-1–11.

[4] A. A. Bell, J. Brauers, and et al., "High dynamic range microscopy for cytopathological cancer diagnosis," *IEEE J. Sel. Topics Signal Process.*, vol. 3(1), 170–184, 2009.

[5] A. C. Ruifrok and D. A. Johnston, "Quantification of histochemical staining by color deconvolution," *Anal Quant Cytol Histol*, vol. 23(4), 291–299, 2001.

[6] C. E. H. Berger, J. A. de Koeijer, W. Glas, and H. T. Madhuizen, "Color separation in forensic image processing," *Journal of Forensic Sciences*, vol. 51(1), 100–102, 2006.