

Segmentation of biological cell images for sonification

Alistair D N Edwards¹, Geneviève Hines¹, Andy Hunt²

¹*Department of Computer Science
University of York
Heslington
York YO10 5DD
UK*

²*Department of Electronics
University of York
Heslington
York YO10 5DD
UK*

alistair@cs.york.ac.uk

Abstract

Cervical cancer is one of the most preventable/treatable forms of cancer, due to the fact that precursor signs for the disease can be detected in microscopic examination of cervical cells. Currently these examinations are purely visual, but the project reported aims to process the visual images to present them in a complementary auditory form and thereby to improve diagnostic accuracy. Standard mean-shift image processing techniques have been successfully applied to extract nuclear data suitable for presentation in an auditory form.

1. Introduction

Cervical cancer is one of the most preventable forms of cancer. This is because it is a slow-onset disease whose precursor signs can be detected by inspecting visually, under magnification, samples of cervical cells. The UK National Health Service (NHS) Cervical Screening Programme organizes the collection and inspection of about 4 million samples each year in England [1]. It is a highly successful program which saves an estimated 4,500 lives each year [*op. cit.*].

While the screening program is evidently highly successful in reducing mortality, it can never be infallible. In particular, since visual inspection of cell samples is a human activity, it is subject to human error. From time to time samples are mis-classified, and both false negative and false positive results have a serious impact. The objective of the project reported herein is to attempt to improve accuracy by giving cytologists an *auditory* representation of the cells which they are inspecting *visually*.

While the emphasis of this project is on the auditory representation of information (sonification), that information is still based on the visual data on the microscope slide. The first phase of this project, therefore, has been concerned with the processing of that image into a form which can then be sonified, and that is the main topic of this paper. The paper presents current work in progress. The image processing techniques that have been applied are not novel or specific to this application, but the purpose of this paper is to present their use in an application which is novel (1) because its ultimate objective is the auditory presentation of information extracted from a visual image and (2) because its application in medical screening is unique.

2. Background

Before presenting the work within this project it is necessary to fill in some background information on cervical screening and sonification.

2.1 Cervical cancer and screening

Cancer Research UK states [2] that “Cervical cancer takes time to develop. There is usually a period when some of the cells lining the cervix develop abnormal changes but are not yet cancerous – these can give rise to cervical cancer later on. Doctors can pick up these changes through screening, and a simple treatment can prevent cancer developing.”

A sample of cervical cells is collected and then either smeared and fixated onto a glass slide (smear method) or preserved in a fluid (Liquid Based Cytology or LBC method) and sent to a laboratory¹.

¹ Traditionally the samples were literally smeared onto the slide and hence the test was often referred to as a ‘smear test’. Now that

At the laboratory, the samples are stained, usually with the Papanicolaou ('Pap') stain. As a result of the staining process, the cells and their major components (cytoplasm, nucleus) are made visible [3].

The slides go through a strict screening process, whose aims are (1) to detect any abnormal cell changes, (2) to assess the type and severity of abnormal cell change, when it is observed and (3) to report the presence of a number of infectious agents, when detected [4]. A sample is called 'negative' when it shows no suspicious nuclear abnormalities and when it is free from such infectious agents.

An adequate LBC slide should contain a minimum of 10,000 to 20,000 squamous cells². As a rule-of-thumb this implies around 20 cells in view at any time using a microscope magnification of 40 times. To screen a single slide manually takes between 10 and 15 minutes, including the paperwork.

Clearly this process is labor-intensive and requires highly skilled operators. There are currently available two commercial automatic screening systems: *Focal Point* and the *Thin Prep Imaging System*, both are FDA³ approved and in use in the US. Two key stages are fully automated: primary screening, where the most normal slides are labeled "negative", and quality control, where a proportion of negative slides are re-selected for a second check. After primary screening, the remaining unlabeled slides must be checked by a human operator, but this can still be aided by the automatic selection of the most suspicious fields of view, so the whole slide need not be inspected [6].

The role currently played by the automatic screening systems is important in the context of the present study because:

- It implies the continuing need for human screeners and hence the need to support them.
- It directs the focus of our sonification work, in that we should not aim to directly detect abnormal cells and then generate some form of alarm sound. Rather we can only present (auditory) information about the cell and leave the screener to interpret that information.

It is also important to emphasize that diagnosis of slides, in particular the grading of abnormal slides, is

not always clear-cut and that opinions sometimes differ, even amongst experienced cytologists.

Cervical cancer is associated with Human Papilloma Virus (HPV) infection. Only a minority of HPV infections will subsequently result in cancer; however nearly *all* women who develop cervical cancer have had past infection with a high-risk strain of HPV. A vaccine to prevent HPV infection (and hence cervical cancer) has now been licensed for use within the European Union. The UK Government is in the process of launching a program of mass vaccination [7].

The advent of vaccination will not obviate the need for cervical testing, though, for a number of reasons. First, while the vaccine prevents the strains of HPV that are most likely to cause cervical cancer, it is not complete protection against all strains. Second, it takes between 10 and 20 years for a cervical cancer to develop after HPV infection and hence the number of women who are unprotected will be significant for some years to come.

2.2. Sonification

"Sonification is defined as *the use of nonspeech audio to convey information*. More specifically, *sonification is the transformation of data relations into perceived relations in an acoustic signal for the purposes of facilitating communication or interpretation*." ([8], original emphasis). An increasing number of examples of sonification exist, but one of the oldest and most familiar is the sonification of radiation levels provided by the Geiger counter. This auditory display has been demonstrated to be superior to a visual equivalent, which is one motivation for the use of sonification [9].

Another motivation is to present additional information during a visually intense task. That information may be redundant in the sense of its duplicating that which is already apparent in the visual image, or it may be additional in representing something which is hard or impossible to detect in the visual image. Even redundant information may have a role to play in that there is evidence that people perform better when receiving complementary information simultaneously on the audio and visual channel [10]. Sonification has already been successfully used in medical applications, including [11].

3. Image processing

We analyze digital images of cervical sample slides to extract the image properties that will drive the sonification system.

liquid-based cytology is often used, it may be inaccurate to refer to a 'smear'. LBC gives better quality slides. The improved quality has made the automatic screening of slides which is discussed below feasible and it is the technique used within the Leeds Trust.

² Squamous cells are not the only ones of interest, but they are larger and easy to count; their density is thus used as a measure of a slide's quality. An LBC slide with fewer than 10,000 squamous cells is likely to be rejected for screening purposes.

³ Food and Drugs Administration

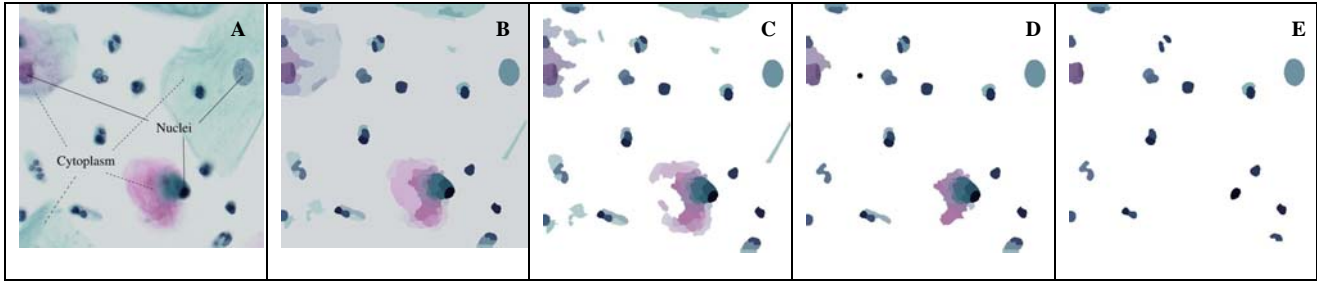


Figure 1. Squamous cells and polymorphs (obj. x40) from an LBC slide (Leeds). Original image and segmentation process (regions are represented by pseudo-colours). A. Original image – B. Initial EDISON segmentation – C. Selection by region size – D. Global gray-level threshold – E. Gray-level threshold on local groups (group definition is based on region connectivity and color similarity).

Cells on a sample slide can occur singly and in different aggregate formations (clumps, balls, etc.) and they tend to form a single layer in LBC slides (less so in a smear). Different cells and cell components take on different colours when stained. While the color intensities and shades are usually well standardized within a single lab, they can vary from lab to lab. A cell consists of a nucleus surrounded by cytoplasm (Figure 1). The aspects of the cytoplasm and the slide background are used by cytologists as contextual information in the assessment of nuclear normality. Features which may be indicative of abnormality include:

- **Size:** In particular an increased relative size of the nucleus compared to the cytoplasm.
- **Shape:** Normal nuclei tend to have a smooth, circular or oval outline.
- **Texture:** Abnormal nuclei may have rough textures (due to chromatin clumping) .
- **Chromaticity:** Abnormal nuclei are often darker (hyperchromatic) and sometimes take a reddish color instead of their usual blue/green

In our prototype system, we first segment the images into nuclei and background (including the cytoplasm and the true image background). We then use texture features to assess the normality of the nuclei because texture on its own is a fairly reliable indicator of nuclear normality. Texture features can also be less sensitive to the outcome of the image segmentation .

For the development of our segmentation algorithm, we used a set of sample slide images captured using a digital camera mounted on the teaching microscope at Leeds Cytopathology Teaching Laboratory. The images (500 x 5000 pixels, magnification x20 or x40) display normal cells and a range of abnormal cells, in single and in aggregate formation. The slides photographed were part of a set of slides used for training purposes and were labelled according to the interesting features of the slide.

Most slides were prepared at the Leeds laboratory, using the *SurePath* Liquid Based Cytology staining technique⁴. A few smears were also used.

3.1. Segmentation

A number of choices were made regarding the desired outcome of the segmentation process:

- It will split the image into one background region and other regions which delineate the cells' nuclei.
- It is allowed to miss very bland nuclei (very likely normal) and out-of-focus nuclei (scanning artefact).
- The nuclear regions can be over-segmented (cut-up into several pieces) because our system will analyze the texture features of groups of connected regions twice: once considering the group as a single region and once considering each region separately.
- The fused/overlapping nuclei can be represented by a single region because fused nuclei are either part of the same cell or are likely to come from cells located nearby on the cervix and therefore are likely to show the same type of abnormality.

The overall scheme for the segmentation process was as follows:

- Perform an initial segmentation of the image where each region contains either nuclear material or background material (including cytoplasm). The regions can be much smaller than a typical nucleus.
- Discard, as far as possible, the background regions.
- Deal with the remaining regions by inspecting groups of connected regions with the aim of discarding remaining background regions and fusing nuclear regions into regions representing, as far as possible, a single, whole, nucleus.

⁴ <http://www.bd.com/tripath/>

It was decided to use the *mean-shift* method, as implemented in the *EDISON* software⁵, to perform the initial segmentation of the slide images. In *EDISON*, this method is supplemented by an edge strength assessment technique. [13-15]. An overview of image segmentation and classification techniques applied to images of tissue section can be found in [16].

The regions most likely to be nuclei are selected on the following criteria:

Region size

Large regions (area over N_{max} pixels) are considered to be background regions.

Gray value

Other regions are first inspected according to a simple grayscale thresholding algorithm: when fewer than $Q\%$ of a region's pixels are darker than a certain threshold, T_g , the region is considered to be a background region.

The reason for this choice is that all nuclei contain chromatin, which is a substance that absorbs the stain more intensely. Chromatin can be present in the nuclei in very fine grains or in clumps, so every nucleus may not be wholly dark, but every nucleus should contain a certain proportion of dark pixels.

The grayscale for each image is linearly stretched to ensure that the range of graylevels present in the image span the full interval (0,255).

Local processing of remaining regions

At this stage in the region selection process, we have one large background region, resulting from the fusion of all the discarded regions and groups of selected (or foreground) connected regions. Each group can be a single nucleus fragmented into several regions, or a nucleus attached to a cytoplasmic region, or a group of nuclei together with some cytoplasm retained between the nuclei.

To further refine the nucleus selection process, groups are further split into subgroups of similar color. Two adjacent regions are placed into different subgroups if the distance between their average colors is more than a threshold T_c and if the two regions cannot be connected by a chain of adjacent regions, where for each link in the chain, the difference between the average color of the two regions is less than T_c . The distance

function used is the Euclidean distance on the 2-dimensional subspace of the $L*u*v$ color space defined by the components $*u$ and $*v$.

A local gray-level threshold T_l is then calculated for each subgroup: the width of valleys in the subgroup's histogram are calculated; among the valleys with a width larger than W , the one corresponding to the brightest gray-levels is chosen and the threshold is placed at its centre. If no valley wide enough is found, the centre of the widest valley is used instead, as long as its width is more than a given W_{min} . If this is not the case, the threshold is placed on the maximum value of the gray-level interval. Then all regions within the subgroup with an average gray level above T_l are discarded as background.

The parameter values chosen to segment our set of images are as follows. For x40 (x20) images, *EDISON* parameters are: *SpatialBandwidth* = 4 (3), *RangeBandwidth* = 60 (20), *MinimumRegionArea* = 150 (100), *GradientWindowRadius* = 4 (2), *MixtureParameter* = .4 (.6), *EdgeStrengthThreshold* = .2 (.4); further parameters are: N_{max} = 2000 (1500), Q = 10 (10), T_g = 175 (175), W = 8 (8), W_{min} = 4 (4).

An example of an unprocessed image and its segmented result is shown in Figure 1.

4. Sonification

Once the nuclei have been successfully segmented out, the aim is to present relevant features in an auditory form. At the current point in this research we have achieved a good enough segmentation to enable us to move on to experiment with the sonification.

We have decided to concentrate on the *texture* of the nucleus. In [19], Walker and Jackaway showed that the statistical geometric features, modified for cytological texture analysis have a high discriminatory power for the two-class problem normal/abnormal cells. We will use the best pair from this set of features to drive our sonification experiments. With two variable features, the texture of any particular nucleus can be represented in a two-dimensional space. The plan is to map out the space of 'typical' abnormal cells and 'typical' normal ones, and then to locate the sound associated with a given cell in that space. In this way the cytologist would hear an audio representation of how well the current cell matched typical normal or abnormal cells. Experiments are under way to determine the auditory dimensions onto which to map these features, based

⁵ *Edge Detection and Image SegmentatiON* system, <http://www.caip.rutgers.edu/riul/research/code/EDISON/index.html>.

on a pilot project we have run as part of this research [20].

Another decision to be made experimentally is which and how many cells to sonify. In order to avoid an auditory averaging effect, our initial choice is to concentrate on sonifying a single cell from each view and that will be the one calculated by the software to be the most 'interesting' cell currently in view. 'Interesting' cells can be the most abnormal cells, however it can also be a class of cells likely to be missed/dismissed, such as "the palest cell showing a given degree of similarity to reference abnormal cells", that is to say within a certain distance of the centre of the abnormal cells distribution.

5. Conclusions

This research is very much still in progress. We have already learned that existing image processing techniques can be applied to the analysis of cell images to extract information which will be capable of being presented in an auditory form. We will shortly be in a position to generate that sonified data and to test its effectiveness. The objective of this phase of the research is to assess empirically whether the cytologists collaborating on our project can indeed benefit from a combined visual/auditory representation of the slides by comparison with a purely visual inspection. If this can be established, then further, more wide-reaching trials will be required to compare the accuracy of slide diagnosis by the current method and by the visual/auditory method, which it is hoped will show that visual+auditory inspection can be more effective.

6. References

- [1] NHS. *Cervical Screening Programme 2005-06 in England*. 2006 [cited; Available from: <http://www.ic.nhs.uk/pubs/csp0506>.
- [2] CRUK. *Cancers at a glance*. 2007 [cited; Available from: <http://info.cancerresearchuk.org/cancerandresearch>.
- [3] Roseblade, R., *Sound and interface design for improving detection of cervical cancer*, in *Department of Electronics*. 2006, University of York: York.
- [4] NHS. *Achievable standards, benchmarks for reporting and criteria for evaluating cervical cytopathology*. 2000 [cited; Available from: <http://www.cancerscreening.nhs.uk/cervical/publications/laboratories.html>.
- [5] NHS, *Laboratory Organisation - A Guide for Laboratories Participating in the NHS*, in *NHS, Cancer screening Programmes*. 2003.
- [6] Hurwitz, H., "The Maturing of the Pap Test". *Physicians Update*, 2006. **XXI**(2): p. 1-4.
- [7] Laurance, J., *Cervical cancer vaccine to be given to all girls aged 12 and 13*, in *The Independent*. 2007: London.
- [8] Kramer, G., et al., *Sonification Report: Status of the Field and Research Agenda*. 1999, National Science Foundation.
- [9] Tzelgov, J., et al., "Radiation detection by ear and by eye". *Human Factors*, 1987. **29**(1): p. 87-98.
- [10] Hapeshi, K. and D. Jones, "Interactive multimedia for instruction: A cognitive analysis of the role of audition and vision". *International Journal of Human-Computer Interaction*, 1992. **4**(1): p. 79-99.
- [11] Fitch, W.T. and G. Kramer, "Sonifying the Body Electric: Superiority of an auditory over a visual display in a complex, multivariate system", in *Auditory display, sonification, audification and auditory interfaces: Proceedings of the First International Conference on Auditory Display (ICAD)*, G. Kramer, Editor. 1992, Santa Fe Institute, Addison-Wesley: Reading, Massachusetts. p. 307-325.
- [12] Penbaker, W.B. and J.L. Mitchell, *JPEG: Still Image Data Compression Standard*. 1992: Kluwer Academic Publishers. 650.
- [13] Meer, P. and B. Georgescu, "Edge Detection with Embedded Confidence". *IEEE Trans. Pattern Anal. Machine Intell.*, 2001. **23** p. 1351-1365.
- [14] Comaniciu, D. and P. Meer, "Mean shift: A robust approach toward feature space analysis". *IEEE Trans. Pattern Anal. Machine Intell.*, 2002. **24** p. 603-619.
- [15] Comaniciu, D. and P. Meer, "Cell image segmentation for diagnostic pathology", in *Advanced Algorithmic Approaches to Medical Image Segmentation: State-Of-The-Art Applications in Cardiology, Neurology, Mammography and Pathology*, J.S. Suris, S.K. Setarehdan, and S. Singh, Editors. 2001. p. 541-558.
- [16] Ong, S.H., X.C.J. Jayasooriah, and R. Sinniah, "Image analysis of tissue sections". *Comput. Biol. Med.*, 1996. **26**(3): p. 269-279.
- [17] Lee, P., *Image (pre)-processing in detecting cervical neoplasia (cervical cancer)*, in *Department of Electronics*. 2006, University of York: York.
- [18] Walker, R.F., *Statistical geometric features - Refinements for cytological texture analysis*. 1996, Department of Electrical and Computer Engineering, University of Queensland: Brisbane. p. 19.
- [19] Walker, R.F. and P. Jackway. *Statistical geometric features - Extensions for cytological texture analysis*. in *Proceedings of the 13th International Conference on Pattern Recognition (ICPR '96)*. 1996.
- [20] Stammers, J., *Developing synthesis techniques for the sonification of precancerous cells*, in *Department of Electronics*. 2006, University of York: York.

Acknowledgement

This work is funded by the UK Engineering and Physical Sciences Research Council, grant number EP/C512413/1. The work would not have been possible without the collaboration of the Leeds

Teaching Hospitals NHS Trust, and particularly
David Brooke.