Preparing Digitized Cervigrams for Colposcopy Research and Education: Determination of Optimal Resolution and Compression Parameters

Jose Jeronimo, MD,[1] Rodney Long,[2] Leif Neve,[2] Daron Ferris, MD,[3] Kenneth Noller, MD,[4] Mark Spitzer, MD,[5] Sunanda Mitra, PhD,[6] Jiangling Guo, PhD,[6] Brian Nutter, PhD,[6] Phil Castle, PhD, MPH,[1] Rolando Herrero, MD,[7] Ana Cecilia Rodriguez, MD,[7] and Mark Schiffman, MD, MPH[1]

[1]*Hormonal and Reproductive Epidemiology Branch, Division of Cancer Epidemiology and Genetics (DCEG), National Cancer Institute, Bethesda, MD;* [2]*Communications Engineering Branch, National Library of Medicine, Bethesda, MD;* [3]*Departments of Family Medicine, Medical College of Georgia, Augusta, GA;* [4]*Department of Obstetrics and Gynecology, Tufts University School of Medicine, Boston, MA;* [5]*Department of Obstetrics and Gynecology, New York Methodist Hospital, Brooklyn, NY;* [6]*Computer Vision and Image Analysis Laboratory, Department of Electrical and Computer Engineering, Texas Tech University, Lubbock, TX;* [7]*Proyecto Epidemiológico Guanacaste, Fundación INCIENSA, San José, Costa Rica*

■ **Abstract**

Objective. Visual assessment of digitized cervigrams through the Internet needs to be optimized. The National Cancer Institute and National Library of Medicine are involved in a large effort to improve colposcopic assessment and, in preparation, are conducting methodologic research.

Materials and Methods. We selected 50 cervigrams with diagnoses ranging from normal to cervical intraepithelial neoplasia 3 or invasive cancer. Those pictures were scanned at 5 resolution levels from 1,550 to 4,000 dots per inch (dpi) and were presented to 4 expert colposcopists to assess image quality. After the ideal resolution level was determined, pictures were compressed at 7 compression ratios from 20:1 to 80:1 to determine the optimal level of compression that permitted full assessment of key visual details.

Results. There were no statistically significant differences between the 3,000 and 4,000 dpi pictures. At 2,000 dpi resolution, only one colposcopist found a slightly statistically significant difference ($p = 0.02$) compared with the gold standard. There was a clear loss of quality of the pictures at 1,660 dpi. At compression ratio 60:1, 3 of 4 evaluators found statistically significant differences when comparing against the gold standard.

Conclusions. Our results suggest that 2,000 dpi is the optimal level for digitizing cervigrams, and the optimal compression ratio is 50:1 using a novel wavelet-based technology. At these parameters, pictures have no significant differences with the gold standard. ■

Key Words: cervigram, digitization, scanning, compression, colposcopy, cervical cancer

Reprint requests to: Jose Jeronimo, MD, Division of Cancer Epidemiology and Genetics, National Cancer Institute, 6120 Executive Boulevard, MSC 7234, Bethesda, MD 20892. E-mail: guibovij@mail.nih.gov

Journal of Lower Genital Tract Disease, Volume 10, Number 1, 2006, 39–44

In virtually all cervical cancer prevention programs, some form of visual assessment is used during the screening or diagnostic process, or both, to determine which women are at risk and require therapeutic

intervention. However, the lack of suitable cervical images covering the full spectrum of cervical neoplasia has hampered the training of colposcopists, assessment of expertise, and research of the natural history of human papillomavirus and resultant cervical neoplasia.

One of our major goals is to help clinicians improve their colposcopic performance. In concert with this goal, the National Cancer Institute (NCI), the National Library of Medicine (NLM), and other collaborators have joined to open access to the many thousands of cervical images taken during NCI-funded research studies (see article in this issue). The cervical images obtained in these studies consist of cervigrams, although colpophotographs are also more commonly obtained in other studies [1].

It has been a major challenge to determine how to incorporate the 50,000 pairs of cervigrams into a database that can be shared easily. Cervicography, described 25 years ago [2], relies on a 35-mm slide (cervigram) taken with a conventional camera equipped with a 100-mm macro lens having a fixed focus. This nondigital photographic technique was used mainly during the last 2 decades. Cervigrams were taken during two NCI-sponsored studies on HPV and cervical neoplasia, the Guanacaste Project [3] and ALTS [4]. Included in these longitudinal series were patients with prevalent invasive cancer or preneoplastic lesions, healthy women at enrollment in whom disease developed during the follow-up, and women who never had pathologic changes in the cervix. This collection gives us a unique opportunity for studying uterine cervix changes related or unrelated to HPV infection, cervical disease, or both.

To manage this extraordinary amount of visual data, we decided to convert the conventional cervigrams into digital files and to compress them to obtain small files that are easy to transmit through the Internet. This article presents our work to determine the optimal resolution and compression parameters for preserving fine visual details of the cervix and acetowhite lesions. The study is of general interest for three reasons: (i) the data demonstrate how important it is to research digital scanning parameters rigorously, because small technical differences can be very important; (ii) the comparison between the experts reviewing identical images reveals the extent to which they agree (or disagree) on even the most basic issue of image adequacy; and (iii) the outcomes will determine optimal scanning and compression parameters for the NCI/American Society of Colposcopy and Cervical Pathology cervical image archive as research begins to explore visual characteristics of colposcopy.

MATERIALS AND METHODS

Cervigrams

We randomly selected from the Guanacaste study [3] a variety of cervigrams demonstrating most of the colposcopic findings of the cervix, including 10 pictures from each one of the following cervicography diagnostic groups: normal, atypical, P0 (probably normal but colposcopy is recommended), P1 (visual equivalent of a low-grade lesion), and P2–P3 (visual equivalent of a high-grade lesion or cancer) [5]. The confidentiality of subjects was strictly maintained, and the cohort study was conducted as an institutional research board-approved research project with written informed consent that specifically mentioned cervicography.

Scanning

The selected 35-mm slides (cervigrams) were scanned using a Super CoolScan 4,000 ED scanner (Nikon Inc, Melville, NY), at 1,500, 1,660, 2,000, 3,000, and 4,000 dots per inch (dpi). Then, one of the evaluators (J.J.) evaluated each one of the cervigrams and marked the area of the cervix with the most visual changes; that area was selected to be presented to the reviewers. The digital files were saved in an uncompressed format, tagged image file format (tiff). The size of the files at each level of resolution is shown in Table 1. Cervigrams scanned at the highest level (4,000 dpi) were used as the gold standard of quality.

Compression

Most conventional compression software packages do not have the flexibility to yield minimum distortion for specific classes of color images. Experts from the NLM and the Computer Vision and Image Analysis Laboratory, Department of Electrical and Computer Engineering, Texas Tech University, are working on alternative methods [6]. Using these methods, the image data are first transformed to the wavelet domain so that one image is represented as a set of wavelet coefficients. Groups

Table 1. Cervical Image Level of Resolution and Size of the Digital File

Resolution (dpi)	Size (MB)
4,000	65.3
3,000	36.7
2,000	16.3
1,660	11.2
1,500	9.2

(or vectors) of coefficients then are represented by numerical codes that are smaller than the original groups. This process forms the basis of vector quantization.

After the evaluators finished the study on scanning features and 2,000 dpi was considered ideal level of resolution, these pictures were selected for the compression study. A small pilot study was undertaken to determine the maximum level of compression to be used in this study (data not shown). In the pilot study, cervical images had an evident loss of quality when compressed at ratios higher than 80:1. Based on these results, for the present study we evaluated 7 levels of compression (20:1, 30:1, 40:1, 50:1, 60:1, 70:1, and 80:1). Uncompressed 2,000-dpi pictures were used as the gold standard.

Cervigrams were compressed in the Computer Vision and Image Analysis Laboratory at Texas Tech University and were sent to the Communications Engineering Branch of the NLM, where the images were processed and adapted to the software developed for this study. The encoding process of an image compression technique can be roughly divided into three major steps: preprocessing, quantization, and postprocessing. Preprocessing generally involves procedures that transform the image from the spatial domain to other mathematical domains to reduce the redundancies and that shifts and scales the pixel values to make them suitable for encoding. Some major preprocessing procedures that occur before vector quantization in HVSQ, a hybrid vector quantization technique developed at Texas Tech University, are color transform, channel mean removal, and discrete wavelet transform. Postprocessing of the vector quantized wavelet coefficients of an image generally includes entropy coding and data packing, so that data can be compressed and organized further. By exploiting the hybrid nature of the algorithm, the HVSQ, modified specifically for uterine cervix image compression from the previous HMVQ compression technique, allows two levels of bit-rate control for each color channel, and hence the technique retains the exact color quality after compression and reconstruction of a medical image. The preservation of color quality of uterine cervix images by HVSQ has been shown to be superior to that of JPEG2000 [7].

Evaluation

Experts from the NLM developed software for comparison of the pictures used in this study. This software, called the Accept program, provided the following functionality: (i) sequential, automatic display of each image in the study; (ii) fast redisplay, or blinking, of overlaid images, under manual control, to allow comparison of test and gold standard images; (iii) automatic recording of observer evaluations to file; and (iv) ability to reevaluate the different resolution or compression representations for a single image.

A screen shot of the Accept software is shown in Figure 1. The software was developed with MATLAB versions 6.5 and 7.1 and was compiled to an executable

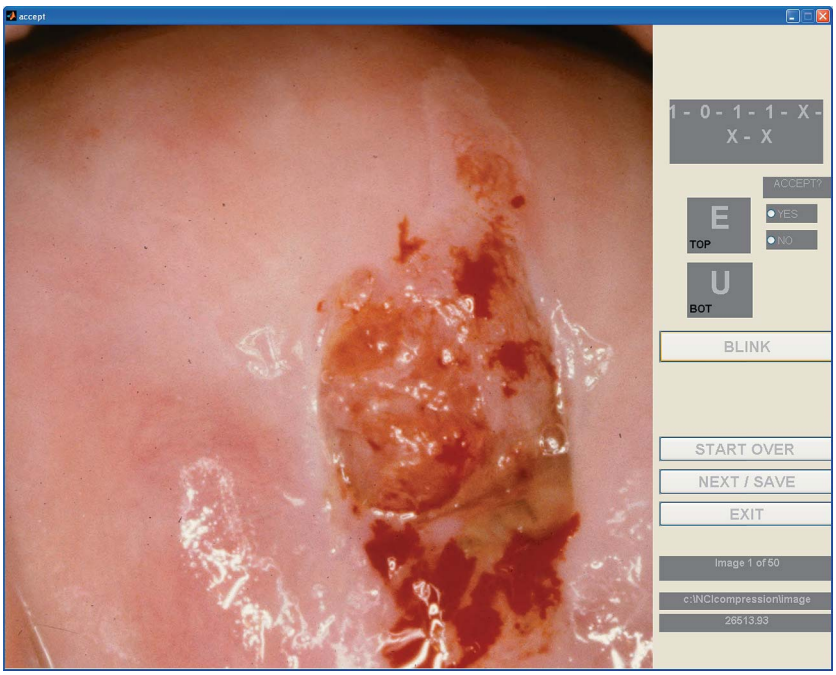

Figure 1. Accept program for comparison of pictures. The fifth compression representation (E) is being displayed. The blink button allows the observer to toggle between the test image (E) and the uncompressed gold standard (U). TOP indicates the image that is currently visible; BOT indicates the image that is not visible. At the top right, Xs serve as placeholders for the observer's evaluations. If the observer accepts a representation, the placeholder for that representation is replaced with a 1; a ''not accept'' becomes a 0. When all representations for an image have been evaluated, the observer proceeds to the next image with the Next/Save button. The Start Over button will clear the evaluations for a single image and allow the observer to begin again for that image. The interface for the resolution study is essentially the same as this interface for the compression study.

version, in which it was distributed to the expert reviewers for execution on Windows PC systems.

In both studies, the displayed images were 1,000 × 1,000 subimages digitally cropped from the original digitized slides. For the resolution study, this enabled the display of the gold standard images on a standard 1,280 × 1,024 display with a direct one-to-one data-to-screen pixel mapping; for the compression study, this enabled the display of both the gold standard and test images with such mapping. For the resolution study, the gold standard images were 1,000 × 1,000 subimages digitally cut from the gold standard 4,000 dpi images. For the test images, which were scanned at 1,500, 1,660, 2,000, and 3,000 dpi, a visual area identical to the gold standard visual area was digitally cut from the original images; then this visual area was scaled by bilinear interpolation to the same 1,000 × 1,000 pixel size. Using this method, the gold standard and test images had the same size on the reviewer's screen.

For the compression study, both the gold standard and test images consisted of 1,000 × 1,000 subimages digitally cut from the original images that had been scanned at the single 2,000 dpi digitization level (as determined by the resolution study). Pertinent software was installed on the personal computers of the reviewers and presented the picture considered as gold standard overlapped and blinking against each one of the images to be evaluated. This process, adapted from similar techniques for detecting astronomical objects against a background of stars, allows the reviewer to focus on localized areas in the images and observe any visual changes between the gold standard and the images evaluated, without changing visual focus.

Pictures were presented in random order. The following anatomical features were considered at the time of evaluating the quality of the picture: squamocolumnar junction; color of acetowhite lesion, if present; borders of acetowhite lesion, if present; and vascular patterns.

Four expert colposcopists (DF, JJ, KN, and MS) were asked to evaluate the quality of the pictures. The reviewers scored the picture as adequate if it was of good quality and preserved the same anatomical details observed in the gold standard picture. Otherwise, the compressed picture was scored as inadequate. Reviewers were unaware of the level of resolution or compression of the pictures (Figure 2).

Statistical Analysis

Lower resolution digitization levels (3,000, 2,000, 1,660, 1,500 dpi) were compared with the matching gold standard 4,000-dpi image and differences in percent image acceptability were tested for statistical significance using the exact McNemar's χ^2 tests. Compressed (at 20:1, 30:1, 40:1, 50:1, 60:1, 70:1, 80:1) 2,000-dpi images were compared with the matching, uncompressed 2,000-dpi image and differences in percent image acceptability were tested for statistical significance using the exact McNemar's χ^2 tests.

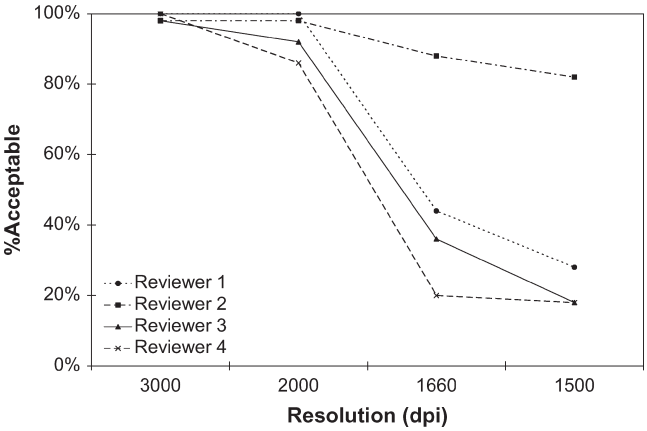

Figure 2. Cervical images considered acceptable per reviewer and level of resolution (n = 50).

RESULTS

To follow the two-step process of managing of visual data, resolution, and compression, this study first determined optimal scanning resolution. Later, after determining the optimal scanning parameters, images using those parameters were used for exploring several levels of compression.

RESOLUTION

Two of the reviewers considered 100% of the 3,000-dpi pictures as adequate for evaluation, but the other two experts found one picture inadequate at this level of resolution (Table 2). No significant differences were noted between these pictures and the gold standard.

When evaluating the pictures with 2,000-dpi resolution, one of the colposcopists considered all of them as having adequate quality. Two evaluators found one and four inadequate pictures, respectively, which were not statistically significant to the gold standard pictures. Meanwhile, one colposcopist scored 7 of the 50 pictures as inadequate at this level of resolution, results that were significantly different (p = .02) to the gold standard.

At 1,660 and 1,500 dpi, all the reviewers found statistically significant difference between the quality of

Table 2. Percentage of Cervical Images Considered Acceptable per Reviewer and Level of Resolution (n = 50)[a]

Reviewer		Resolution (dpi)			
		3,000	2,000	1,660	1,500
1	Acceptable n (%)	50 (100)	50 (100)	22 (44)	14 (28)
	p^b =	n/a	n/a	<0.00005	<0.00005
2	Acceptable n (%)	49 (98)	49 (98)	44 (88)	41 (82)
	p =	1.0	1.0	0.03	0.004
3	Acceptable n (%)	49 (98)	46 (92)	18 (36)	9 (18)
	p =	1.0	0.1	<0.00005	<0.00005
4	Acceptable n (%)	50 (100)	43 (86)	10 (20)	9 (18)
	p =	n/a	0.02	<0.00005	<0.00005

n/a, not available.
[a]Cervical images scanned at 4,000 dpi were considered the gold standard.
[b]McNemar $\chi 2$ test between the current resolution level and the gold standard.

these pictures and the gold standard. There was a clear loss of resolution quality of the pictures between 2,000 and 1,660 dpi, which was evident in the clinical (picture 2) and statistical evaluation (Table 2). We concluded that 2,000 dpi represented a reasonable tradeoff of file size and resolution quality.

COMPRESSION

In the second part of the study, pictures at 2,000 dpi were compressed and compared against the uncompressed pictures. The compression ratios were 20:1, 30:1, 40:1, 50:1, 60:1, 70:1, and 80:1. As shown in Table 3, one of the reviewers found a statistically significant loss of quality of pictures compressed at ratio 30:1 or higher; but the other three evaluators did not find that difference. A second evaluator considered that, at a compression ratio of 40:1, there was significant difference compared with the gold standard, but this same evaluator did not find that difference at compression ratio 50:1.

At compression ratio 60:1, 3 of 4 reviewers found statistically significant differences when comparing these pictures against the gold standard; and at ratio 70:1, all the colposcopists unanimously considered them to be inadequate.

DISCUSSION

When digitally scanning photographic images, one must select the optimal resolution to satisfy the intended purpose. Selection of a suboptimal resolution will yield lower quality digitized images. Conversely, selecting an excessive level of resolution generates enormous file sizes and superfluous visual information.

This study was developed with the objective of determining the most effective way to digitize cervigrams for use in a very large research and education program accessible by experts with different internet access and computer capabilities. Consequently, we needed to determine a minimal level of cervical image resolution sufficient to preserve important visual details while simultaneously minimizing resolution to easily handle

Table 3. Percentage of Cervical Images Considered Acceptable per Reviewer and Ratio of Compression (n = 50)

Reviewer		Compression ratio						
		20:1	30:1	40:1	50:1	60:1	70:1	80:1
1	Acceptable n (%)	50 (100)	50 (100)	50 (100)	47 (94)	33 (66)	19 (38)	10 (20)
	p =	n/a	n/a	n/a	0.3	<0.00005	<0.00005	<0.00005
2	Acceptable n (%)	48 (96)	48 (96)	46 (92)	48 (96)	44 (88)	43 (86)	36 (72)
	p =	0.2	0.5	0.1	0.5	0.03	0.02	0.0001
3	Acceptable n (%)	50 (100)	49 (98)	50 (100)	49 (98)	46 (92)	44 (88)	39 (78)
	p =	n/a	1.0	n/a	1.0	.1	.03	.001
4	Acceptable n (%)	48 (96)	44 (88)	44 (88)	44 (88)	36 (72)	21 (42)	15 (30)
	p =	0.5	0.03	0.03	0.03	0.0001	<0.00005	<0.00005

n/a = not available.
[a]Uncompressed cervical images were considered the gold standard.
[b]McNemar $\chi 2$ test between the current compression level and the gold standard.

file transfers. The same premises applied to the compression process.

Our cervical image resolution results suggest that 2,000 dpi is the optimal level for digitizing cervigrams. Greater levels of resolution would produce much larger file sizes. Put in perspective, in the two NCI studies, approximately 100,000 cervigrams were taken. If 2,000-dpi digital files (16.3 MB each picture) were used, we would need to dedicate 1.63 terabytes (1,630 GB) of our server for archiving that visual data. Any increase in the size of the files would have a large impact on the space needed for storage.

Even though the optimal scanning features give us high-quality digital pictures preserving all visual details of the cervix, the size of the files makes them impractical for access through the Internet. We are using novel compression software that, unlike most of the commercial compression software available, is based on wavelet technology that allows us to reach a high compression level with minimal loss. Using this software, we can create images almost identical to the uncompressed file at compression ratios as high as 50:1.

This study has determined the optimal levels of resolution and compression of 35-mm cervical images. We are now completing digitization of the cervigrams and refining the several database tools that will access them (see article in this issue). We look forward to sharing the NCI/American Society of Colposcopy and Cervical Pathology cervical image archive in the future with researchers around the world, in our joint efforts to improve understanding of the healthy and neoplastic cervix. There are several studies to be developed using the described digitized cervigrams such as: number of human papillomavirus (HPV) types infecting the cervix and number of acetowhite lesions; visual patterns according to HPV type; spectrum of lesion characteristics in high-grade squamous intraepithelial lesion, low-grade squamous intraepithelial lesion, atypical squamous cells of undetermined significance, and ASCH patients; colposcopic patterns of women with positive HPV results; colposcopic patterns of women with negative HPV results; visual appearance according to viral load; patterns of appearance and disappearance of acetowhite lesions; and so forth.

REFERENCES

1. Ferris DG, Willner WA, Ho JJ. Colpophotography systems: a review. *J Fam Pract* 1991;33:633–9.

2. Stafl A. Cervicography: a new method for cervical cancer detection. *Am J Obstet Gynecol* 1981;139:815–25.

3. Herrero R, Schiffman MH, Bratti C, Hildesheim A, Balmaceda I, Sherman ME. Design and methods of a population-based natural history study of cervical neoplasia in a rural province of Costa Rica: the Guanacaste Project. *Rev Panam Salud Publica* 1997;1:362–75.

4. Schiffman M, Adrianza ME. ASCUS-LSIL Triage Study. Design, methods and characteristics of trial participants. *Acta Cytol* 2000;44:726–42.

5. Schneider DL, Herrero R, Bratti C, Greenberg MD, Hildesheim A, Sherman ME. Cervicography screening for cervical cancer among 8460 women in a high-risk population. *Am J Obstet Gynecol* 1999;180:290–8.

6. Mitra S, Yang S, Kustov V. Wavelet-based vector quantization for high-fidelity compression and fast transmission of medical images. *J Digit Imaging* 1998; 11:24–30.

7. Guo J. SPKKYSMSNB. Bit-rate Allocation Control and Quality Improvement for Color Channels in HMVQ Image Compression. IEEE Symp. on CBMS, 110–5. Bethesda: IEEE; 2005.