1-21-2009

Mark Detection from Scanned Ballots

Elisa H. Barney Smith
Boise State University

George Nagy
Rensselaer Polytechnic Institute

Daniel Lopresti
Lehigh University

Mark Detection from Scanned Ballots

Elisa H. Barney Smith[1], George Nagy[2], Daniel Lopresti[3]
[1]Boise State University, [2]Rensselaer Polytechnic Institute, [3]Lehigh University
EBarneySmith@boisestate.edu, nagy@ecse.rpi.edu, lopresti@cse.lehigh.edu

ABSTRACT

Analyzing paper-based election ballots requires finding all marks added to the base ballot. The position, size, shape, rotation and shade of these marks are not known a priori. Scanned ballot images have additional differences from the base ballot due to scanner noise. Different image processing techniques are evaluated to see under what conditions they are able to detect what sorts of marks. Basing mark detection on the difference of raw images was found to be much more sensitive to the mark darkness. Converting the raw images to foreground and background and then removing the form produced better results.

Keywords: mark detection, ballot image processing, forms processing, background removal

1.0 INTRODUCTION

One of the foundational elements of a democracy is the election. Countries around the world handle the technical implementation of the election differently. Voting methods currently in use in the United States can be classified as (1) direct reading electronic (DRE) (2) punchcard, (3) lever, (4) paper for automated reading ("op-scan") and (5) paper for human counting. Voting machines with levers that turn mechanical gear counters are being abandoned because of the prohibitive cost of replacement parts and because they cannot satisfy federal handicapped voter requirements. Punched card readers can fail to count partially removed chips, like the hanging chads in Florida that made headlines in the year 2000 US presidential elections. DRE requires software that is distrusted by many voters because of the lack of direct evidence that their vote is recorded and counted as cast. Human counting is too slow. Optical scan therefore is a mode being increasingly used.

In its draft report on Voluntary Voting Systems Guidelines for 2007 [2], the Security and Transparency Subcommittee (STS) for the Technical Guidelines Development Committee of the National Institute of Standards and Technology (NIST) wrote that there are open technical issues that can and should be addressed:

> "The widespread adoption of voting systems incorporating paper did not seem to cause any widespread problems in the November 2006 elections. But, the use of paper in elections places more stress on (1) the capabilities of voting system technology, (2) of voters to verify their accuracy, and (3) of election workers to securely handle the ballots and accurately count them. Clearly, the needs of voters and election officials need to be addressed with improved and new technology. The STS believes that current paper-based approaches can be improved to be significantly more usable to voters and election officials…"

In response to this need, we have started to explore some of the document analysis issues related to the op-scan technology. This draws on techniques developed in both the document recognition and retrieval communities.

Optical scan ballots occur in many formats. Voters can fill ovals or rectangles, or connect lines. These are intentional *marks* designed to be counted as votes. It is possible that other marks not intended as votes get added either by the voter, or through handling. The voter then either runs the ballot through a scanner (sometimes called a Portable Ballot Counter or PBC), or puts it in an envelope for processing at election headquarters after voting closes.

While most ballots come with explicit instructions such as "completely fill in the oval(s) next to your choice(s)," a non-negligible percentage of the voters do not heed these instructions. The legal definition of a valid vote in most US districts is linked with "voter intent." Therefore to count all votes any mark that has been added to the ballot should be identified and analyzed, be it in the specified location and the specified shape or not. This bears a similarity to form processing

where the user entered data is extracted from the pre-printed form and is analyzed. Forms processing often makes use of operator confirmation. It can also exploit database constraints, and is primarily an effort to determine the text in each field (OCR). For ballot reading while the majority of the marks will be in the prespecified target positions with a known shape, the marks can be in any position on the ballot and can take any shape. This changes the problem from one of detecting the presence or absence of ink in a specific position on a form, to one of locating and classifying all inputs to a form that were made by the voter. We believe that detailed characterization of the marks on each ballot may provide politically acceptable criteria for objective determination of valid votes.

Although optical mark sensing has only recently become a popular voting option among concerned citizens and legislators, this technology dates from the 1950s. Jones [4] provided a solid summary of the existing automatic ballot reading technologies. Many of the older readers only check for marks of sufficient size in discrete regions of the ballot. Newer ballot counting devices employ a CCD desktop scanner technology, as it is well developed and moderately priced, but they still only check discrete regions of the ballot. This technology will be assumed and used in the following experiments and discussions as methods to detect any shape or size mark are evaluated. The results will to a large degree transfer to images acquired with other scanning technologies.

Since there are already commercially successful op-scan election products in the marketplace, the first goals of our research are not to propose our own competing approach for reading ballot markings. Rather, we are working to bring rigorous analyses of the problems that arise in the processing of scanned ballots to the scientific literature. Given the fundamental role of elections in democratic societies, news stories concerning voting issues frequently appear in the popular media. There is, however, no comparable body of published work that examines the important questions at a technical level. While voting system vendors may have their own internal (closely-guarded) quality control processes, the widely publicized flaws that have already been identified by independent security experts demand a more open approach. It is our hope to bring members of the international document analysis research community into this discussion.

The dataset and some of the data quality issues are described in Section 2. Methods to do form removal are compared for printed and scanned ballots in Section 3. The performance on various mark styles and qualities are compared and analyzed in Sections 4 & 5.

2.0 THE DATASET

Op scan ballots display instructions and candidate information printed in dark ink on a light colored solid background. To this the voter is expected to fill in an oval or square, or connect two parts of a broken arrow to indicate his/her preference. The ballot images used in this experiment are fill-in-the-oval style ballots. They were marked using the PERFECT Ballot generation software [5]. This gave precise control over the shape, shade, size, rotation and position of the markings. The base ballot template is from Minnesota. A sample is shown in Figure 1a. It nominally supports 17 different races with up to five candidates, one allowing for votes for two candidates. Five sets of ballots were marked. To more fully utilize the form contents and reduce processing time, all of the 60 targets were marked. While we desire to develop algorithms that can detect marks anywhere on the ballot, initially marks in or near the targets are generated, but the position of the marks is not used in the detection algorithms. The marks are made in the shapes of x's, checks, solid ovals or small dots, and in this dataset were produced with frequencies of 30%, 22%, 25% and 23% respectively. The small dots could represent votes or hesitation marks. In this paper we are just aiming to evaluate the effectiveness of image processing methods to identify all the marks on the ballot with a limited number of false alarms for a range of scanning configurations and marks. In a future step the marks will then be classified as to noise, unintentional voter marks, or intentional voter mark styles. The positions of the marks were varied so there was a range of overlap with the target ovals. The marks were printed in a range of gray levels and a range of relative sizes. The gray levels represent variability in marking instruments. Samples of marks are shown in Figure 1b.

The synthetic ballots were printed on a standard office inkjet printer with no intentional degradations. The forms are nominally printed at the same size and scanned at the same resolution, but humidity can change the size of the document by as much as 1% which can be $1/10^{th}$ of an inch [4]. Indeed, failure to accommodate this led to SAT scoring problems in October 2005 [3]. For this dataset, the five marked and one blank synthetic ballot were then scanned in gray level at 200dpi. The lighting will be assumed to be fairly uniform. In a real ballot reading scenario the template and the marked ballot can both be scanned at a variety of conditions. While the Ballot Counting Devices should have high quality optical scanners and should be tuned in advance to produce a high contrast output, most ballot readers are used once a year, are

stored in a warehouse when not in use and are operated by volunteer election staff, or in some cases by the general voting populace themselves. This means that to have a robust system, a wider range of inputs needs to be explored. The data set used here was designed to reflect particularly poor scanning quality to more thoroughly test the image processing algorithms.

Each printer and scanner will have its own set of characteristics so the relative gray levels are not guaranteed to be maintained. Two primary types of input scanning were used in these experiments. One was designed to be "lighter" than the other, resulting in a background paper color that was nominally white, while the other "darker" resulted in a background that was a light gray. The images were created with known set of gray levels. The form text itself was a dark gray close to black, and the paper was initially pure white. The field headers contained regions of a solid light gray. Four shades of grey were used for the markings. Only a small set of the marks were created in a pure black shade (5%). The images were then printed on a high quality inkjet printer and scanned. The gray values used in the source image, and the gray values that resulted after printing and scanning are shown in Table 1, along with the frequency of each base gray level. Most of the marks were a nominal "full" size, but a few were deliberately made at 50%, 75%, 125% and 150% of this base size.
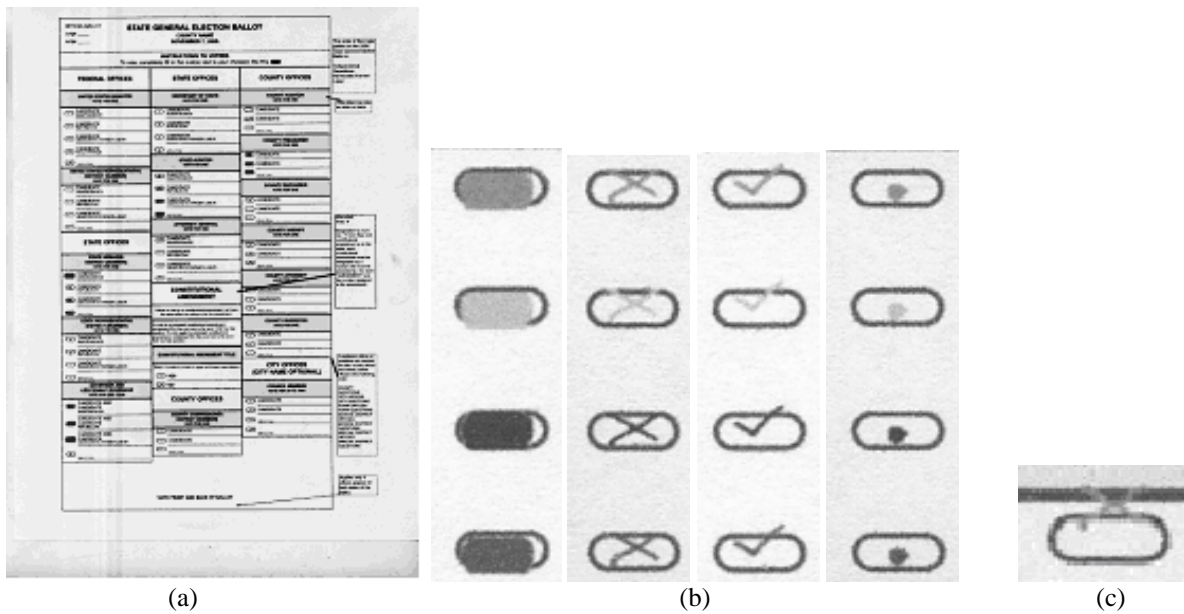

| (a) | (b) | (c) |

Figure 1: (a) Sample ballot image. (b) Close-up view of sample marks with a range of gray levels. (c) Example of a mark that will be difficult to detect with almost any method.

Table 1: The gray values in a [0,255] scale of the initial image components and the resulting shades after printing and scanning.

Initial Image Shade	Content	Number of marks per dataset	Resulting Image Shade Scanner setting "dark"	Resulting Image Shade Scanner setting "light"
0	Marks	15	81	65
32	Form lines		86	71
64	Marks	15	99	83
128	Marks	195	142	143
192	Marks	75	193	199
231	Header Zones		224	233
255	Paper		249	255

3.0 PROCESSING

To distinguish user marks from the preprinted ballot contents, the known ballot contents need to be identified and removed from the marked ballot image. The method used in this paper to do this is to register the template image to the marked ballot image and to compare these images. This is similar to form dropout used in form processing [7,16]. It also has relationships to foreground/background separation used in motion detection [10]. Success in this process is the identification of all user-added marks. Several methods used for comparing the template and the marked form are explored.

The printing and scanning processes introduce two levels of degradation and uncertainty to the image content. The first is the intensity noise and gray level mappings. The second is the change in position, scale and skew of the ballot image. In order to detect all marks made on a ballot beyond what was present in the template, image registration and image subtraction form the basis of the processing. Various methods to pre- and post-process the images and the image differences were explored to see which would make the marks detectable while reducing the number of false alarm marks that would have to be processed in follow-on work to decide if they were marks or noise.

The first step in the processing was image registration. A rough estimate of the relative angle, scale and translation was attained through using a Fourier Mellin phased-based correlation technique [11], for scale and skew, followed by regular phase correlation for translation. The correlation was performed on down-sampled versions of the images for speed and memory efficiency. A single target (oval) example was manually extracted from the ballot template. All occurrences of this target example were found in the full resolution template image. It was skewed and scaled per the initial estimate and the skewed and scaled version was found in the marked ballot. These mark positions produced the set of control points used in final image registration. To determine control point correspondence, the initial estimate of the relative document skew, scale and translation was used to transform the control point positions in the template image. The sets of control points were then compared to find those that were in close geographic proximity based on the coarse transformation estimate to distinguish erroneous detections from true marks. This set formed the set of corresponding control points that were used to infer a linear conformal spatial transformation. This spatial transformation was then applied to the template image to align it with the marked image. The transformed template was compared to the scanned marked ballot to determine possible marks. For most pairs the registration process produced very close alignment, however, even with most of the 60 targets being identified and matched, the estimate of skew and scale in one of the ten images maintained enough error that the registration mismatch over the whole image was visible. The average rotation estimated was 0.25^o with a maximum of 0.61^o. That is still enough that the estimation and correction are vital. The scale was always estimated to be very close to 1. The translation ranged 1 to 16 pixels, rarely in integer amounts.

A range of different image processing techniques were explored to see which would be most robust to the matching and best enable extraction of the added marks while reducing the number of false alarm marks detected. Further processing at a later step could be used to distinguish true marks from false alarm marks, but it is desired to initially have the fewest false alarms possible.

The difference images will ideally display just the image components that were added to the marked ballot. There are two types of noise that can be present to inhibit this. One comes from the misregistration, which includes the errors that occur from trying to rotate, scale, or translate images on a square grid by non square units. The other comes from the differences between the two images resulting from scanning noise or unequal presence of paper edges. The boundaries of the initial ballot template paper were recorded. These coordinates were transformed along with the ballot template image. Regions outside these boundaries are known to be scanner background and not paper, and thus can not contain any voter added marks. These regions were masked out and not included in mark detection.

In [1] five image preprocessing methods were explored. Those methods are summarized in Table 2. The first four consist of various combinations of absolute difference between the target and the registered reference image, with or without smoothing before or after the difference. The differenced image is then thresholded to decide whether the difference constitutes a mark. These points are then closed with a disk structuring element of radius 5 to merge pieces broken across strokes, or through thresholding after convolution. After thresholding the connected components were identified. Components smaller than 2x2 were discarded. The rest were evaluated for being true marks or false alarms. These same processing methods form the basis of the experiments in this paper on scanned ballot images.

Table 2: The five basic image preprocessing methods used

Method	Description
D1	The absolute value of the difference between the raw images.
D2	The absolute value of the difference between the raw images, smoothed by a 3x3 uniform kernel.
D3	The absolute value of the difference between 3x3 smoothed images.
D4	The absolute value of the difference between 3x3 smoothed images, smoothed by a 3x3 uniform kernel.
D5	The absolute difference of smoothed adaptively thresholded images.

4.0 EXPERIMENTS

Exploration of ballot mark detection was previously done on synthetic ballots that had not been printed or scanned so they were noise-free [1]. The differencing methods summarized in Table 1 showed their potential given a range of mark sizes, shades and shape, as well as looking at the effect of when the mark is in the clear or when it overlaps some pre-printed feature of the ballot. The methods are evaluated here again with the synthetic images that will later be printed and scanned to provide a baseline for the optimal case. Then the images produced when these are printed and scanned are used for another set of experiments to show the challenges introduced by uncontrolled scanning.

In [1] it was found that when thresholding the difference images to separate marks from non-marks higher thresholds will detect fewer marks and lower thresholds detect more marks. It was found that the adaptive threshold was necessary to detect the lightest of the marks, at a relatively high threshold level, but that it introduced other artifacts which caused a higher false alarm rate. Those same experiments were run on the synthetic source images produced for this dataset, before they were printed and scanned. This will allow for better comparison with the results of the scanned ballots. The ballots are at half the resolution of the previously analyzed ballots. The results are summarized in Table 3. As before a lower threshold was needed to detect the lighter marks. For method D1 all marks, including the low contrast ones, could be perfectly detected if the proper threshold was used. The smoothing in methods D2-D4 meant that an even lower threshold was needed as the thin strokes when blurred had a lower maximum intensity difference from the paper. Also it occasionally led to broken strokes after thresholding as the stroke widths and neighborhoods did not contain constant amounts of "ink", some of which was fixed through the application of a closing operation. False alarms were almost exclusive to the adaptive threshold method D5, because it produced artifacts when the neighborhoods varied in content, different thresholded images were produced even if the source image was identical except for the mark.

Mark detection after printing and scanning is a more realistic problem and is more difficult due to the added noise from printing, scanning and registration. The methods D1-D4 were also applied to the printed and scanned images in the dataset in this current paper. Samples of the resulting difference images for the four methods for light and dark scanning are shown in Figure 2. Due to the gray scale mapping, different thresholds are needed for these images than were appropriate for the synthetic images. Also since there are two sets of scanner settings, the threshold that works on one scanner setting likely will not work on images produced with the other scanner setting. These same difference image segments are shown in Figure 3 after thresholding at level 60.

Performance of mark detection can be evaluated in a similar framework to information retrieval through the use of the metrics precision and recall metrics,

$$\text{Precision} = \frac{N_{\text{Detected}}}{N_{\text{Detected}} + N_{\text{FA}}} \tag{1}$$

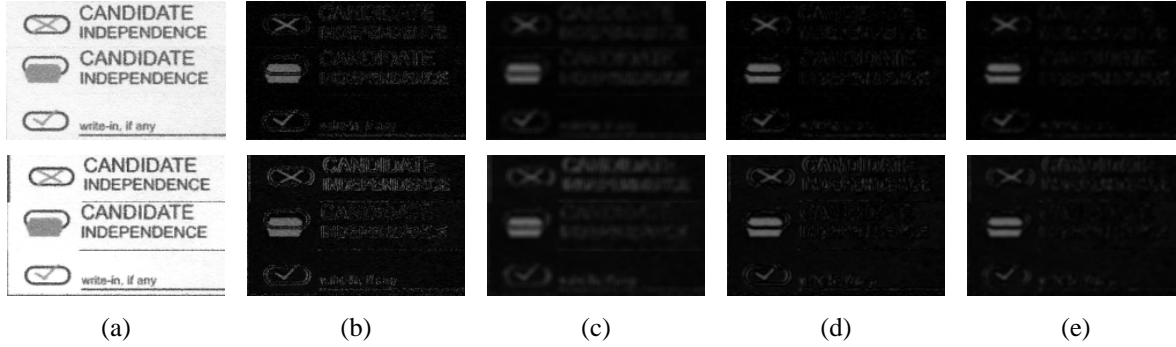$$\text{Recall} = \frac{N_{\text{Detected}}}{N_{\text{True}}}, \tag{2}$$

Figure 2: (a) samples from original image, (b)-(e) corresponding samples when differenced with methods D1-D4. Top row is for dark scanned original, bottom row is for light scanned original. Black is no image difference.
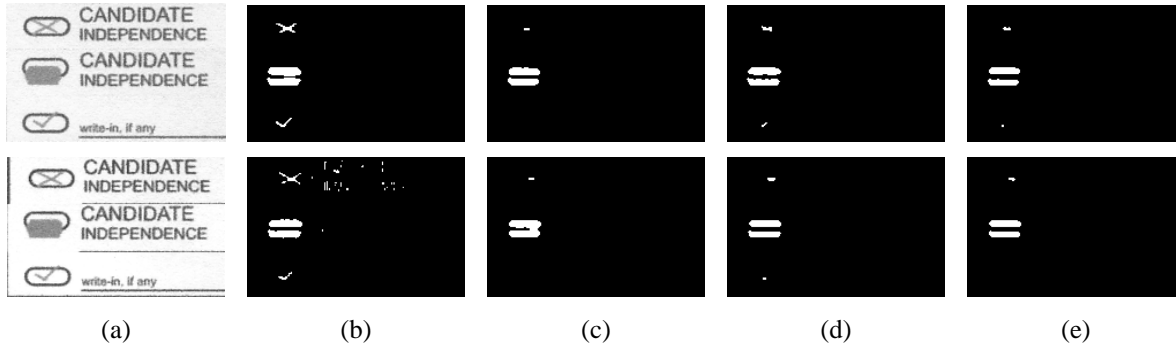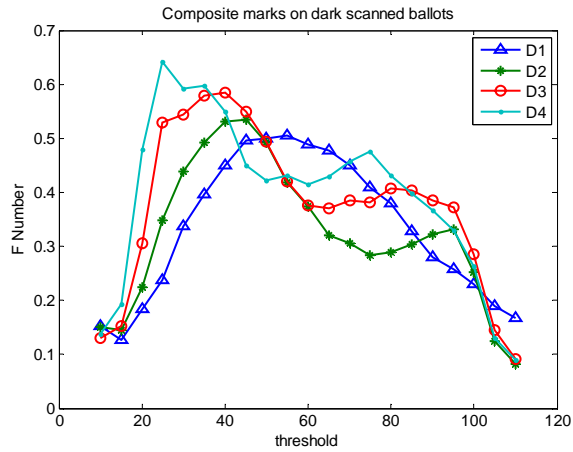

Figure 3: (a) samples from original image, (b)-(e) corresponding samples when differenced with methods D1-D4 and thresholded at level 60. Top row shows results for dark scanned original, bottom row shows results for light scanned original. White is detected mark.

where N_{Detected} is the number of marks found that correspond to true marks, N_{FA} is the number of marks found that do not correspond to true marks and N_{True} is the number of true marks available to be found [6]. The F-measure which is a single measure that trades off precision and recall is the harmonic mean of precision and recall

$$F = \frac{2PR}{P+R}.$$ (3)

The F-measure is shown for each of the four different methods D1-D4 over a range of thresholds in Figure 4. No one method performs uniformly best and the 'optimal' threshold depends on the differencing method. The noise present in the image causes the intensity to spread to adjacent bins. This leads to more false alarms, and the difference in scanning leads to variations in the threshold's effectiveness. At low thresholds image components can merge resulting in a loss of detection. Table 4 shows the results with three global threshold settings. They were chosen based on the gray levels shown in Table 1 to somewhat correspond to the thresholds used in the synthetic experiments. As a comparison, a threshold was chosen to be halfway between the maximum and the minimum image difference. This was more robust than a fixed threshold, but still missed several marks.

To try and mitigate the variation in intensities, the scanned images could be thresholded to separate foreground marks and form elements from background paper and noise. The difference between the thresholded image and the template image would then show the additions on a high contrast image. This somewhat reverses the steps used previously where the difference was thresholded. Common methods to do thresholding include global and adaptive thresholds. First the global threshold determined by Otsu [9] was applied. The fact that the low contrast marks are such a small fraction of the image area makes Otsu less effective for this problem, as often the marks are removed by the thresholding process so they were not present when compared to the unmarked image (Table 5). A different method to select a global threshold was developed to take advantage of the a priori information about the ballot structure containing large areas of header zone gray that is lighter than the lightest added mark. A threshold was chosen at the valley in the intensity histogram of
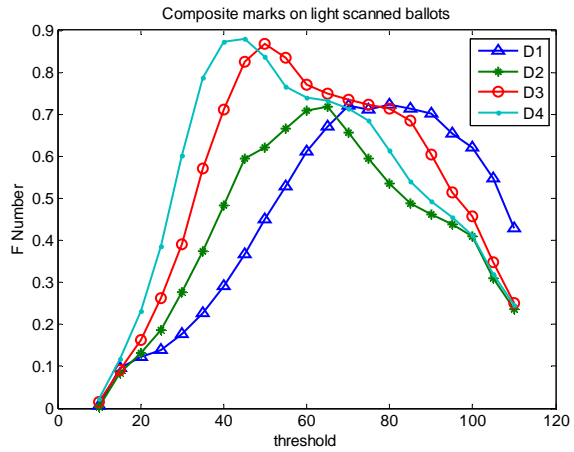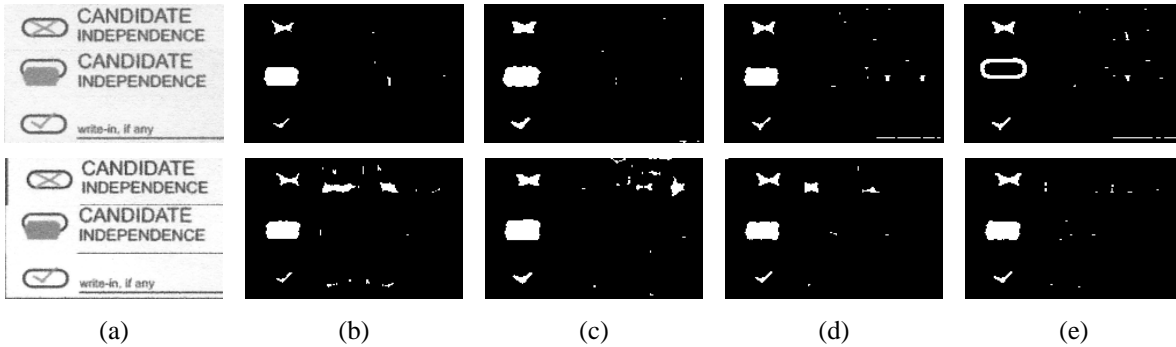
Figure 4: F-measure plots for the four differencing methods, (a) for dark scanned ballots (b) for light scanned ballots.

Figure 5: (a) samples from original image, (b) difference of images binarized by Otsu's global threshold (c) difference of images binarized by a trough based global threshold (d) difference of images binarized by Niblack's adaptive threshold (e) difference of images binarized by Tsinghua's adaptive threshold. Top row is for dark scanned original, bottom row is for light scanned original.
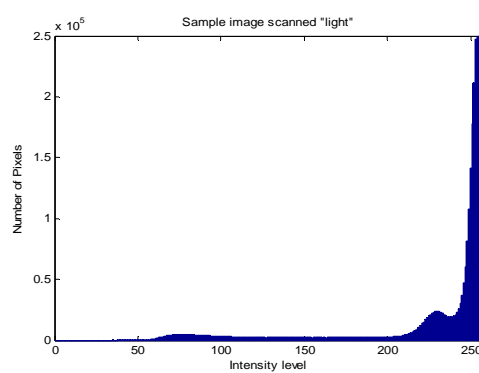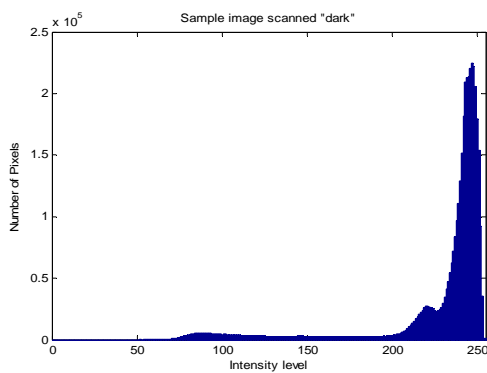

Figure 6: Histogram of gray levels present in (a) one "dark" and (b) one "light" image. Count for intensity level 255 in (b) has been cropped for display. It reaches 1.5×10^6.

the trough between the brightest gray level peaks, Figure 6. If the gray zones were not present, an alternative method would be to pick a threshold just as the peak corresponding to the paper tapers off. Here better performance was attained (Table 5), but the variability in the solid gray regions due to the tail being split led to an increase in false alarms.

An alternative to global thresholding is adaptive thresholding as in method D5 from [1]. Two adaptive methods were explored for front-end processing, the Niblack method [8] with a window size of 15 and k=-0.2 and one developed by Tsinghua University [15] with a window size of 25 and mean C=0.03. These adaptive methods have the advantage that the low contrast marks should be retained with less dependence on the contrast. They do require a post-processing step to remove "ghosts" [13] which is computationally expensive. The parameters for these two adaptive methods were not fine-tuned for this dataset. It is expected that variable scanning conditions will always be present and tuning for ballots read at each election precinct is not a practical step. The results of these methods are also shown in Table 5 with examples in Figure 5. While more of the marks were detected, the false alarm rate (Table 6) is often increased. The false alarm rate is shown in Table 6. The images scanned darker were more prone to false alarms than those scanned lighter. The dark scanning made more image and noise features visible. Also the one ballot that did not align well was included in the dark scan data results. The Tsinghua adaptive thresholding method seemed to produce the highest detection rate with a relatively lower false alarm rate. Otsu had low false alarms, but also low detection.

Table 3: Mark detection results for synthetic images

Differencing Method	Threshold		
	50	60	175
D1	100%	99%	30%
D2	95%	91%	13%
D3	95%	95%	14%
D4	95%	92%	14%
D5	100%	99%	81%

Table 4: Mark detection results for scanned images

Differencing Method	Scanner Configuration	Threshold			
		80	90	150	Auto
D1	Dark	58%	49%	3%	84%
D2	Dark	37%	28%	1%	56%
D3	Dark	48%	36%	2%	76%
D4	Dark	38%	29%	1%	65%
D1	Light	61%	55%	9%	84%
D2	Light	37%	30%	7%	46%
D3	Light	55%	43%	8%	67%
D4	Light	44%	33%	7%	49%

Table 5: Mark detection results for thresholded scanned images

Differencing Method	Scanner Configuration	Thresholding Algorithm			
		Otsu	Trough	Adaptive T	Adaptive N
D1	Dark	69%	74%	89%	89%
D2	Dark	56%	80%	83%	76%
D3	Dark	67%	82%	89%	88%
D4	Dark	64%	81%	87%	84%
D1	Light	81%	81%	94%	91%
D2	Light	58%	84%	84%	75%
D3	Light	75%	88%	93%	89%
D4	Light	67%	87%	90%	84%

Table 6: Average number of false alarms per page for thresholded scanned images

Differencing Method	Scanner Configuration	Thresholding Algorithm			
		Otsu	Trough	Adaptive T	Adaptive N
D1	Dark	105	166	126	224
D2	Dark	94	114	90	126
D3	Dark	123	111	107	186
D4	Dark	80	99	72	123
D1	Light	82	164	102	184
D2	Light	6	39	< 1	24
D3	Light	8	35	< 1	39
D4	Light	7	23	< 1	30

5.0 CONCLUSION

An evaluation of several methods of detecting marks in printed ballots has been conducted. The dataset focused on low contrast marks where the majority were not solid filled ovals. Thus the optimal conditions generally expected for op-scan ballot readers were not present. It has shown that detection is in general possible, although perfect detection was not achieved.

There were a fair number of false positive marks found in the data analyzed. Some of these were caused by marks detected on the page border where the mask did not reach. Others were from misalignment and still others were from inherent differences in the sample ballot pages. In Ballot Image Processing, the cost of a miss is much higher than the cost of a false alarm. All marks greater than 2x2 pixels were considered marks. One possibility to improve the final results is to focus on detecting all the true positives, and then to apply some post processing to analyze the size and shapes of the marks detected to classify them as true positives or false positives. The marks in this dataset were quite diverse, but in a real situation, a voter will usually use the same marking instrument (pen or pencil) for the entire ballot. Also a voter will be more inclined to either always fill in the ovals solidly or always express his or her candidate preference with an "X". A comparison of each detected mark candidate with the average or median of all the marks on the same ballot could provide more information. If most of the marks are large checkmarks, then it would be reasonable to classify small marks of other shapes as noise or as hesitation marks. The mathematical framework for this kind of analysis, dubbed style, appears in [12,14].

This work does not define the end result of a perfect ballot reading algorithm, but does show some of the processing conditions and the results that are possible under them to let the public and vendors know the limits of the optical ballot reading process. In future work, other background removal algorithms will be explored, as will post-processing algorithms to reduce the false positive rate.

6.0 ACKNOWLEDGMENTS

7.0 REFERENCES

[1] Elisa H. Barney Smith, Daniel Lopresti, George Nagy, "Ballot Mark Detection," Proc. International Conference on Pattern Recognition 2008 (ICPR 08), Tampa, Florida, USA, 8-11December 2008.

[2] William Burr, John Kelsey, Rene Peralta, John Wack, "Requiring Software Independence in VVSG 2007: STS Recommendations for the TGDC," National Institute of Standards and Technology, November 2006. http://vote.nist.gov/DraftWhitePaperOnSIinVVSG2007-20061120.pdf (accessed March 21, 2008).

[3] College Board Releases SAT Answer Sheet Processing Report, http://www.collegeboard.com/press/releases/110853.html , 20 July 2006.

[4] Douglas Jones, "Counting Mark-Sense Ballots, Relating Technology, the Law and Common Sense," Part of the Voting and Elections web pages, http://www.cs.uiowa.edu/~jones/voting/optical/ , 2002.

[5] Daniel Lopresti, George Nagy, and Elisa H. Barney Smith, "A Document Analysis System for Supporting Electronic Voting Research," Proc. Document Analysis Systems, Nara, Japan, September 2008, pp. 167-174.

[6] Christopher D. Manning, Prabhakar Raghvan and Hinrich Schütze, "Introduction to Information Retrieval," Cambridge University Press, 2008.

[7] Jianchang Mao, K. Mohiuddin, "Form Dropout Using Distance Transformation," Proc. Conference on Image Processing, 23-26 October 1995, Washington, DC, USA, pp. 328-331.

[8] Wayne Niblack, "An Introduction to Digital Image Processing," Prentice Hall, 1986, pp. 115-116.

[9] Nobuyuki Otsu, "A threshold selection method from gray-level histograms," IEEE Trans. Systems, Man and Cybernetics, Vol. 9, No. 1, January 1979, pp. 62-66.

[10] Massimo Piccardi, "Background subtraction techniques: A review," Proceedings IEEE International Conference on Systems, Man and Cybernetics, 10-13 October 2004, The Hague, Netherlands, pp. 3099-3104.

[11] B.Srinivasa Reddy and B.N. Chatterji. "An FFT-based Technique for Translation, Rotation, and Scale-invariant Image Registration," IEEE Trans. Image Processing, Vol. 3, August 1996, pp. 1266-1270.

[12] Prateek Sarkar, George Nagy, "Style consistent classification of isogenous patterns," IEEE Trans. PAMI, Vol. 27, No. 1, January 2005, pp. 88-98.

[13] Øivind Due Trier, Torfinn Taxt, "Evaluation of Binarization Methods for Document Images," IEEE Trans. PAMI, Vol. 17, No. 3, March 1995, pp. 312-315.

[14] Sriharsha Veeramachaneni, George Nagy, "Analytical results on style-constrained Bayesian classification of pattern fields," IEEE Trans. PAMI, Vol. 29, No. 7, July 2007, pp. 1280-1285.

[15] Guanglei Xiong, "Local Adaptive Thresholding," Mathworks File Exchange, October 2005, http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=8647 .

[16] Bin Yu, Anil K. Jain, "A Generic System for Form Dropout," IEEE Trans. PAMI, Vol. 18, No. 11, November 1996, pp. 1127-1134.