9-18-2011

Evaluation of Voting with Form Dropout Techniques for Ballot Vote Counting

Elisa H. Barney Smith
Boise State University

Shatakshi Goyal
Boise State University

Robbie Scott
Boise State University

Daniel Lopresti
Lehigh University

DOI: 10.1109/ICDAR.2011.101

Evaluation of Voting with Form Dropout Techniques for Ballot Vote Counting

Elisa H. Barney Smith
ECE Department
Boise State University
Boise, ID, USA
EBarneySmith@BoiseState.edu

Shatakshi Goyal
ECE Department
Boise State University
Boise, ID, USA

Robbie Scott
ECE Department
Boise State University
Boise, ID, USA

Daniel Lopresti
CSE Department
Lehigh University
Bethlehem, PA, USA
lopresti@cse.lehigh.edu

***Abstract*—Vote counting accuracy has become a well-known issue in the vote collection process. Digital image processing techniques can be incorporated in the analysis of printed election ballots. Current image processing techniques in the vote collection process are heavily dependent on the anticipated, geometric positioning of the vote. These techniques don't account for markings made outside of the requested field of input. Using various form dropout techniques, however, every mark on the form can be extracted and used by the machine to make an intelligent decision. Most methods will still miss a few marks and result in a few false alarms. This paper explores methods of voting between the results of the different mark extraction methods to improve recognition. To provide diversity a simple image subtraction technique is paired with a distance transform and a morphology based algorithm. The result has a higher detection rate and a lower false alarm rate.**

***Keywords*- mark detection, combination techniques, form dropout.**

I. Introduction

The use of paper ballots can provide independent auditing capabilities to elections, however they do not, guarantee fair, verifiable, fast and efficient elections. The Security and Transparency Subcommittee for the Technical Guidelines Development Committee of the National Institute of Standards and Technology (NIST) observes in a draft report on Voluntary Voting Systems Guidelines for 2007 [1] that the use of paper to provide independent auditing capabilities in elections is entirely practical, but that there are undeniably open technical issues that can and should be addressed.

Early ballot-readers used optical mark recognition (OMR) and mark-sense readers with discrete photocells. Here marks were expected and sensed solely in discrete locations, and their presence determined by the response in those discrete regions exceeding a threshold. Although virtually all ballot readers now use optical scanners with CCD or CMOS arrays, many ballot scanners still simply mimic OMR.

Work by Nagy et al. [2] has found ways to reliably determine the locations of the target ovals assuming the form was designed for reading on a fixed grid OMR system. This can find the ovals and the marks if they can be identified, but often the voter will erroneously choose to mark the form in locations other than in the ovals and knowledge of the target locations will not help in finding these marks.

While the instructions provided to voters are usually a variant on "To vote, completely fill in the oval(s) next to your choice(s) like this: ", in the United States the legal definition of a vote is based on voter intent and thus marks that do not precisely follow the instructions, either by varying the mark shape, fill ratio, marking instrument darkness, or mark position may still be valid and need to be identified and analyzed.

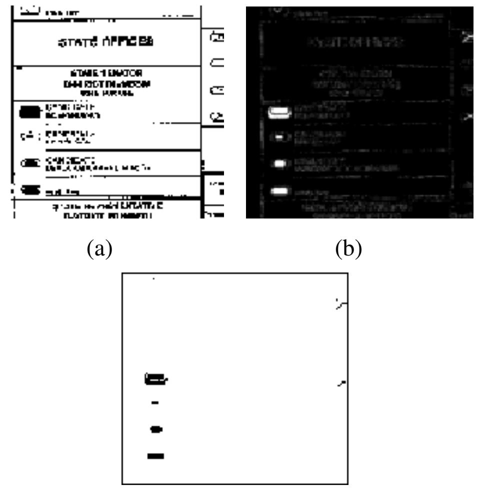
In [3] and [4] some simple methods of locating marks on ballots were explored, first on synthetic ballots, then on ballots printed and scanned. Four

DOI: 10.1109/ICDAR.2011.101

mark detection methods involving subtraction of a blank ballot image from a filled-in ballot image were explored. As expected, not all marks were detected, and marks categorized as false alarms were detected; at lower thresholds more marks were detected but so were more false alarms, and the opposite at higher thresholds. Ballots that were scanned at a lower contrast/brightness setting, thus appearing 'darker' had more image noise to contend with than those scanned at a brighter or 'lighter' setting. The best results achieved for detection was 89% on the 'dark' ballots and 94% on the 'light' ballots, however the associated false alarm rates were 107 and 102 per page, respectively.

This paper explores the benefit of running multiple mark detection methods and then *voting* on the results to increase detection rates while decreasing false alarm rates. Three methods of form removal have been explored. The first is the best performing Image Subtraction method from the earlier work [4], the second is based on a Distance Transform [5] and the third is based on a Morphological Transform [6].

Section II describes the three form removal methods that are being applied to the ballot reading problem. Section III describes the voting procedures. Experiments and results to test their efficacy are described in Section IV. The conclusions are portrayed in Section V.



(a) (b) (c)

Figure 1. (a) Original image (b) Difference image before thresholding, (c) Difference image after thresholding.

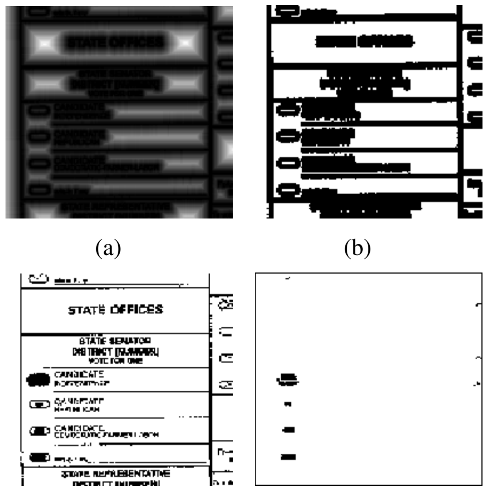
II. Form Removal Methods

The preprinted form information can, in theory, be removed from a filled in form by aligning the two images and performing an image difference. Problems result from imperfections in the image alignment and variations in the image printing and sampling which cause the template form and the filled-in forms to differ by more than just the added marks. Three methods that address this are described next.

A. Image Subraction

In prior work [3][4] four methods of taking the difference between a template and a marked ballot image were presented. These methods alternately took the raw images, or a smoothed version of the image, then calculated the absolute difference and either returned that raw difference or the smoothed version of it. It was determined that if the ballot images came from gray level scans, that performance was greater if prior to differencing, the images were thresholded, rather than taking the difference of gray scale images and thresholding the result. For this paper, the Otsu threshold is applied. Depending on the thresholding technique, the recognition rate ranged from 56% to 94%, and the average number of false alarms per page ranged from less than 1 to 186. As expected, the higher detection rates were usually accompanied by the higher false alarm rates. The best performing of the 4 variations came from taking the convolution of both the thresholded template image and the thresholded filled image, then taking the absolute difference, and smoothing that result to reduce noise, Figure 1b. The smoothing in both cases was done with a 3x3 uniform kernel. This difference image is then thresholded again to determine what content is considered as a mark and what content is considered as noise, Figure 1c.

DOI: 10.1109/ICDAR.2011.101



(a) (b)

(c) (d)

Figure 2. (a) Example Distance Transform. Lighter values indicate higher distance. (b) Unsafe regions shown in black. (c) Original marked ballot. (d) Extracted marks.

B. Distance Transform

The distance transform form removal method [5] is based on evaluating how far pixels in the image are from known text or rulings on the template form. To compensate for the inexact match between the edges in the aligned forms, safe and unsafe regions are defined in the scanned template. The scanned gray scale image is converted to black and white by using a global threshold (Otsu), as in Method 1, and from this the distance transform (DT) is calculated [7]. The DT calculates the distance from each pixel to the nearest black pixel, Figure 2a. Pixels that are within a certain distance from the black pixels in the blank form are considered ’unsafe,’ while those further away are considered ‘safe,’ Figure 2b.

This algorithm extrapolates from the pixels on the border between the safe and unsafe regions to find the rest of the filled-in data. Using information gained from taking the gradient of the DT image, the un-safe region is searched in the gradient direction for black pixels which are used to redraw the user entered data where previously a form line was drawn. Figure 2d shows a portion of the final image after this algorithm is complete.

C. Morphological Subtraction

The Morphological Transform technique [6] was originally developed to extract handwritten information on bank checks. It was assumed that the filled check content differed from the blank only in specific areas and that the difference would be the data entered by the user. It also assumed that there could be some difference between the template and the filled check images, and this method was designed to take these minor variations into account.

The first step in the morphological transform method is to subtract the template from the filled-in form, Figure 3a,b. An erode operation with a 3x3 ‘plus’ shaped structuring element (SE) is applied to get rid of the pepper noise, Figure 3c.

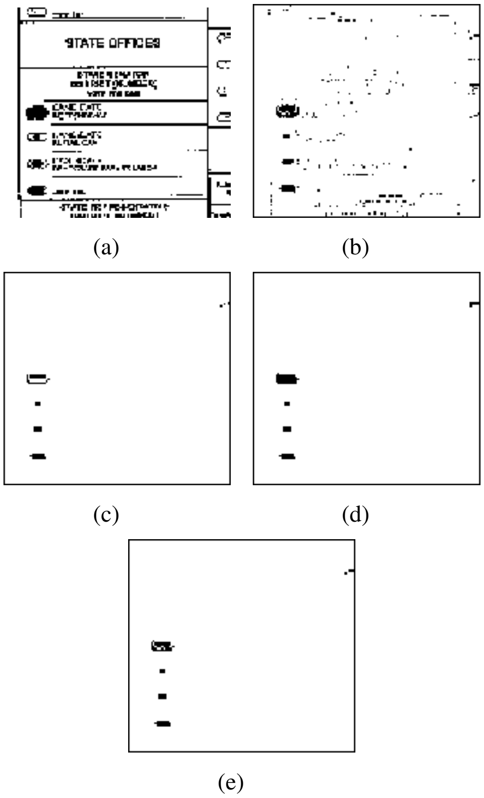
To remove any additional noise and to help further clean up the image a closing operation is performed with a circular SE with diameter 26, Figure 3d, which is followed by subtracting the filled-in form from the current image. Finally a logical AND operation is performed between the image resulting and the simple subtraction image to obtain the final image, Figure 3e.

III. VOTING

In OCR the results of post processing techniques can be used to reduce the error rate, but in mark detection that same procedure can not be used because the location and contents of the voter’s additions is unknown. In theory it will be limited to filling in targets reducing the problem to an OMR problem, but in reality that is not the case. Using an ensemble of classifiers for a pattern recognition problem can improve the performance of the total system [8]. The main issues are to choose the ensemble classifiers and to determine how to combine the results. Ideally the classifiers will have uncorrelated errors. This is accomplished through diversity. In the mark detection application, similar requirements exist. The form removal methods described in Section II were chosen because they have different approaches and thus increase the diversity.

The combination of the results takes a different form for mark detection than it does for general

DOI: 10.1109/ICDAR.2011.101

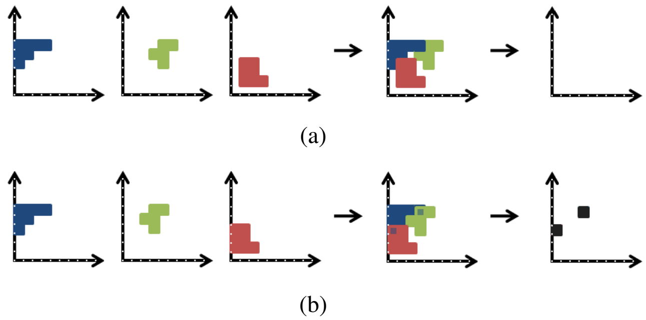


(a) (b) (c) (d) (e)

Figure 3. Examples from Morphological Form Removal. (a) Original and (b) raw difference image. Result of (c) erode with a '+', (d) close with circle. (e) Final Image.

pattern recognition in that the number of classes is no longer restricted to a small discrete set. The detected mark size, shape and position can vary considerably, and interpreting whether the marks detected by competing algorithms match or not involves more than checking whether the class is an 'a', 'b' or 'c'. If all the marks detected by each method had the same shape, size and relative positions, but a random subset of the union of all marks was chosen by each algorithm, then the voting would resemble the voting in traditional pattern recognition. In mark detection, the marks have different shapes and sizes in addition to a 'random' decision to be detected or not. Thus when trying to decide if a mark has been detected by multiple algorithms, how to count the multiple detections becomes an issue.

A variety of methods were considered for the voting method. Voting methods included looking



(a)

(b)

Figure 4. If the same mark is detected, but in offset positions, there may be (a) no detection (b) or multiple detections after voting.

at whether the detected marks overlap in a 'raw' state such that pixels were labeled as marks by two of the three methods. As the shape of the detected marks could be irregular due to border noise, it is possible that different parts of the same mark would be detected by the different detection algorithms. In that case the raw 'on' pixels would not contribute to the vote correctly causing a miss, or they could create multiple disjoint detected marks in a close vicinity of each other, Figure 4. Therefore different voting methods were explored to see which would give better net results.

A raw count of which pixels in an image were classified by 2 or more of the three form removal methods was the base voting scheme. A valid mark (whether a true vote or a false alarm) was defined in [3] and [4] to be a connected component of size greater than 2x2 pixels. This was defined to reduce false alarms from small noise occurrences. The mark shape that remains after raw summing of the images could also have connected components less than 2x2 in size, even if marks were found in the same relative location on the ballot. Voting both with the original image and with the small connected components removed was evaluated.

To reduce the effects of edge noise on the detected marks, each of the results from Methods 1-3 was then dilated and then votes on them were taken. Structuring elements of square sizes 3x3 and 5x5 were used. This produced 8 voting options.

A ninth voting method considered whether the

DOI: 10.1109/ICDAR.2011.101

center of masses of the connected components detected by multiple methods were within a specified distance. Defining the centers of the connected component as the average of the extents produced the tenth voting method.

IV. Experiments and Results

Experiments were conducted using form images generated by the BallotTool [9]. Five pages of ballot images were created. While the goal is to be able to identify marks added by a voter in any location on the ballot, not just in the target regions, the marks in these ballot samples are limited to the target regions, which only affects the experiments in increasing the proximity of form content to the added marks. Each page had 60 marks, with a mixture of filled ovals, check marks, x-marks and dots. All of these marks have been observed in the Minnesota 2008 senatorial race ballot set [10]. The positioning of these marks relative to the target center varied, as did their size, darkness (gray level) and rotation angle. These were printed and scanned twice each, once with a high brightness setting, and once with a low brightness setting, forming two subsets 'light' and 'dark'. The blank ballot template page was also scanned at both brightness settings. The scanned area that lies outside the paper boundaries was removed.

The template image and the test image were aligned by roughly estimating the relative angle, scale and translation with down-sampled versions of the images using a Fourier Mellin phase-based correlation technique [11] for scale and skew. This was followed by extracting the target oval positions in both the template and the marked ballot images, using the initial angle, scale and translation estimate to get mark correspondence and estimating the linear conformal spatial transformation. This was applied to the template image to align it with the marked image as to not further distort the markings made by the voter.

To create bilevel images, a global threshold with the Otsu threshold was applied to both the blank form template and the marked ballots. The three form removal methods were each applied to the 10 form images.

The 10 variations on voting methods described in Section III were applied to the three outcomes for each form. An example of the marks detected before and after voting is shown in Figure 5. Of the 10 voting schemes tried, the best performance was from taking the raw detected mark, prior to 2x2 size limiting, dilating that by a 5x5 square SE, and forming the final image based on pixels now turned 'on' in two or more of the three form removal algorithms.

Initially the three mark detection algorithms were tuned to be rather conservative on detection to reduce the false alarm rate significantly. The detection rate decreased after voting in many of these cases because at some pixel locations only one mark detection algorithm detected a mark, while in other places three algorithms detected marks as desired. With the false alarm rate being relatively small it was not noticeably affected. Because the sensitivity to the false alarms would be reduced if a voting scheme were implemented, the final threshold in Method 1 was lowered to increase its detection rate.

Table I shows the detection and false alarm results for five different thresholds in Method 1. Performance of mark detection is presented as percentage of the marks detected (% det) and average number of false alarms per page (# FA). Voting does not always increase the detection rate and decrease the false alarm rate at a specific threshold, but for a given false alarm rate detection increases, and for a given detection rate, false alarms decrease.

V. Conclusions

The goal of this project is to produce field data, after form dropout, which could be successfully interpreted by a mark detection process. After implementing these three methods of form removal, we are trying to make a comparison to decide which is most likely to work for a wide range of possible mark styles for election ballots. Each individual method was able to detect many marks, but also with false alarms. The combination of the

DOI: 10.1109/ICDAR.2011.101

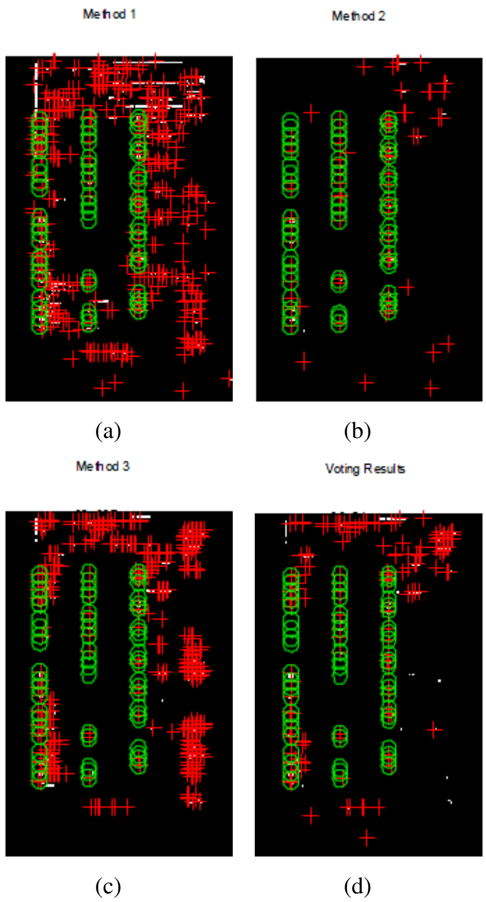


Figure 5. Individual results for 3 form dropout methods plus the result after voting. Target locations are shown in green. Red indicates positions of detected marks.

three methods was able to produce a combined result that was better in both detection and false alarm avoidance, not just moving to another point along the precision/recall curve. Further gains may be possible by varying the sensitivity of the morphological and distance transform methods to increase the range of inputs to the voting process.

Mark detection is a new application in the document image processing field. The voting method presented has applications to the form reading fields from which methods have been taken to enhance this problem. In addition to helping increase accuracy of mark detection, good voting methods could help in more traditional document image analysis techniques such as determining whether to consider extracted data as text or extraneous noise when reading forms, increasing accuracy in line removal applications, or maybe even simple image binarization.

Table I

RESULTS ON ‘DARK’ AND ‘LIGHT’ SCANNED DATA. DETECTION RATES AND FALSE ALARMS ARE SHOWN FOR A RANGE OF THRESHOLDS FOR METHOD 1 ALONG WITH THE RESULTS FOR METHODS 2 AND 3. THE VOTING RESULTS AFTER VOTING ARE SHOWN AT THE SAME THRESHOLDS.

	Dark						
M1 thr	64	75	85	102	128	M2 Dist	M3 Morph
% Det	93%	93%	89%	85%	66%	84%	50%
# FA	96.0	39.0	1.8	0.4	0	1.4	0
Voting							
% Det	92%	92%	92%	91%	85%		
# FA	2.2	1.4	0.6	0.2	0.2		
	Light						
M1 thr	64	75	85	102	128	M2 Dist	M3 Morph
% Det	73%	72%	70%	67%	60%	69%	52%
# FA	136.4	52.4	6	0.4	0.2	1.2	0
Voting							
% Det	72%	72%	72%	71%	68%		
# FA	1.8	0.4	0.2	0.2	0.2		

REFERENCES

[1] W. Burr, J. Kelsey, R. Peralta, and J. Wack, “Requiring software independence in VVSG 2007: STS recommendations for the TGDC,” Technical report, National Institute of Standards and Technology, November 2006.

[2] George Nagy, Daniel Lopresti, Elisa H. Barney Smith, Ziyan Wu, “Characterizing Challenged Minnesota Ballots”, Procs. Document Recognition and Retrieval, San Jose, SPIE, Vol. 7874, January 2011.

[3] Elisa H. Barney Smith, Daniel Lopresti, George Nagy, “Ballot Mark Detection,” Proc. International Conference on Pattern Recognition XIX, Tampa, FL, December 2008.

[4] Elisa H. Barney Smith, George Nagy, Daniel Lopresti, “Mark Detection from Scanned Ballots,” Procs. Document Recognition and Retrieval, San Jose, SPIE, Vol. 7242, January 2009.

DOI: 10.1109/ICDAR.2011.101

[5] Jianchang Mao and K. Mohiuddin, “Form Dropout Using Distance Transformation,” Proc. Int. Conf. on Pattern Recognition, 1995, pp. 328-331.

[6] Yimei Ding, Furnitaka Kimura, Minoru Okada, Malayappan Shridharb and John W. V. Miller, “Background removal for check processing using morphology,” SPIE Two- and three-dimensional systems for inspection, control, and metrology II, Vol. 5606, 2004, pp. 19-26.

[7] Gunilla Borgefors, “Distance Transformations in Digital Images,” Computer Vision, Graphics and Image Processing, Vol 34, 1986, pp. 344-371.

[8] L. I. Kuncheva, Combining Pattern Classifiers, Wiley, New Jersey, USA, 2004.

[9] Daniel Lopresti, George Nagy, Elisa H. Barney Smith, “A Document Analysis System for Supporting Electronic Voting Research,” Procs. Document Analysis Systems, Nara, Japan, September 2008, pp. 167-174.

[10] Elisa H. Barney Smith, Daniel Lopresti, George Nagy, “Towards Improved Paper-based Election Technology,” Proc. International Conference on Document Analysis and Recogntion, Beijing, China, September 2011.

[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.