



# A Document Analysis System for Supporting Electronic Voting Research

Daniel Lopresti\*  
lopresti@cse.lehigh.edu

George Nagy†  
nagy@ecse.rpi.edu

Elisa Barney Smith‡  
EBarneySmith@boisestate.edu

## Abstract

*As a result of well-publicized security concerns with direct recording electronic (DRE) voting, there is a growing call for systems that employ some form of paper artifact to provide a verifiable physical record of a voter's choices. In this paper, we present a system we are developing to support a multi-institution, cross-disciplinary research project examining issues that arise when paper ballots are used in elections. We survey the motivating factors behind our work, discuss the special constraints raised in processing ballots as opposed to more general document images, and describe the current status of our system.*

## 1. Introduction

Florida's infamous "butterfly ballot" and its "hanging chads," resulting in politically-charged calls for a recount during the 2000 U.S. Presidential Election, have initiated the most dramatic series of changes in America's voting history. Soon after the subsequent push for widespread adoption of electronic voting equipment, computer security experts and concerned citizens began raising serious questions about the reliability and trustworthiness of such systems when collecting, storing, and tabulating votes [3, 6, 7, 8, 14, 23]. Direct Record Electronic (DRE) voting, once seen as a straightforward albeit expensive solution, is increasingly viewed as an unacceptable compromise [5, 20]. As a result, there is a growing call for voting systems that employ some form of paper artifact to provide a verifiable physical record of a voter's choices. Often, this takes the form of a hand- or machine-marked paper ballot which is processed by an optical scanning system and then safely secured in the event a recount becomes necessary.

Paper is not new to elections, of course, and issues relating to the design and use of paper ballots have been exten-

sively studied in the past [9, 15]. Mark-sense readers preceded OCR by several decades: the IBM 805 Test Scoring Machine was introduced in 1937. Because punch card technology was already well developed, it was easy to add sensors to detect conductive marks made by "electrographic" pencils. Optical Mark Recognition (OMR), introduced in the 1950's, measured reflectivity instead of resistance, allowing use of a much wider range of writing instruments for multiple choice examinations. For standardized examinations, the questions were printed next to the "bubbles." For low volume tests, general-purpose answer sheets, with numbered rows and columns of bubbles, were used in conjunction with an instruction sheet. The bubbles were relatively large and the marks had to fill most of the bubble in order to be surely sensed. This technology was eventually adapted to election ballots modeled on the classic Australian secret ballots which had been in use since 1858 (when they were, of course, counted by hand).

The development of reliable, low-cost facsimile scanners had an impact on OMR just as it did on OCR. The sensors now swept the entire page, and even small marks were covered several times by several sensors as the paper transport slid the page under the sensor bar. Image processing techniques could now be applied to discriminate marks according to shape and color rather than only total reflectance. This led to new designs for more space-efficient and more easily applicable marks, such as horizontal or vertical line segments in a diamond, or filling a gap in a broken arrow.

Despite this history, introducing (or, rather, re-introducing) paper into the modern election process is not without its own controversy. While there are a myriad of non-technical concerns that have been voiced, ranging from a feeling that paper is "old fashioned," to worries about increased "litter," and the added cost of the required consumables (paper ballots, ink/toner for printers, etc.), eliminating or mitigating drawbacks associated with hardcopy offers an intriguing set of problems that merit serious consideration by the document analysis research community.

In a draft report on Voluntary Voting Systems Guidelines for 2007 [4], the Security and Transparency Subcommittee for the Technical Guidelines Development Committee of the National Institute of Standards and Technology

---

\*Department of Computer Science and Engineering, Lehigh University, Bethlehem, PA 18015, USA.

†Department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY, 12180, USA.

‡Department of Electrical and Computer Engineering, Boise State University, Boise, ID 83725, USA.

(NIST) observes 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:

“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 ...”

Today, paper is treated as an “add-on” when it should be an integral component in trustworthy voting technology. Much attention is now being paid to other parts of election systems; substantial benefits should likewise accrue when paper artifacts are accorded the attention they deserve.

Examples of the issues faced in using paper abound. In Chester County, Pennsylvania, a close election that would determine the majority party in the State House of Representatives was disputed when one party insisted that a recount be conducted by running the optical scan ballots through a different brand of scanner hardware, noting that the tallies can vary depending on the system in use [1]. Interestingly, the same recommendation arose in consulting firm Booz Allen Hamilton’s analysis of a well-publicized problem in scoring the October 2005 SAT test [2]. An apparent discrepancy between paper ballots that had been machine-counted versus those that had been hand-counted led to a heated debate in the 2008 New Hampshire Democratic Primary [24].

There is, of course, a connection between ballot reading and automatic forms processing, a topic which has been heavily studied in our field (*e.g.*, [11, 12, 13]), as well as to the scoring of standardized tests, as noted earlier. Processing paper ballots used in elections differs from these other tasks in important ways, however. The range of individuals who use a paper ballot is likely to be much greater than other forms applications, since all citizens meeting certain basic requirements are entitled to vote in a country’s elections. A certain percentage of voters are only semi-literate, non-native speakers, or suffer from various disabilities that may interfere with their ability to read or mark a ballot. Another requirement which can be legally mandated is that ballots must preserve a voter’s anonymity. This precludes including any unique identifier on the ballot in advance of the election, as well as attempts to contact a voter after-the-

fact should his/her selections prove unreadable. Since elections are held infrequently, voting equipment sits unused for months-on-end, often in storage environments that are not conducive to longevity of the hardware. Maintaining chain-of-custody is a critical security requirement for all election records. Finally, while there is no direct financial interest in an election’s results, there is tremendous public interest; the process of casting and counting votes must be transparent and trustworthy.

In this paper, we present a software system we are developing to support a multi-institution, cross-disciplinary research project examining issues that arise when paper ballots are used in elections. Our *BallotTool* application consists of a collection of software components integrated in a graphical user interface with versions that run under the Linux and Microsoft Windows operating systems. Functionality is provided for a number of document analysis and human factors experiments we are planning to conduct.

While our examples are drawn from recent experiences in the United States, many of the questions we are studying have implications in other countries as well; with the instantaneous communication provided by the Internet, a voting system proven insecure in one locale cannot be trusted in any other. We survey the motivating factors behind our project, discuss the special constraints raised in processing ballots as opposed to more general document images, and describe the current status of our system. Further details concerning our ongoing work can be found on the PERFECT project website [22].

## 2. Issues in Electronic Voting Research

The system we are building consists of a number of software components for supporting voting research centered around paper ballots. This includes tools for defining ballot layouts, processing voted ballots, conducting recounts, and confirming the accuracy of machine ballot reading. We are also developing frameworks for conducting large-scale studies of human and computer performance in interpreting the marks made by voters on ballots. Below, we briefly highlight some of the basic problems we are working to address in our research:

- *Undetected failures in the machine reading of ballots.* As noted previously, there is usually no warning when recognition errors arise in optical scan systems [1, 2]; processing the ballot a second time may lead to a different result [25].
- *Systematic errors due to ballot layout.* Our past work in OCR demonstrated that recognition errors are not uniformly distributed across the page [17]; the same observation may be true of ballots, a fact which may disadvantage one candidate over another based purely

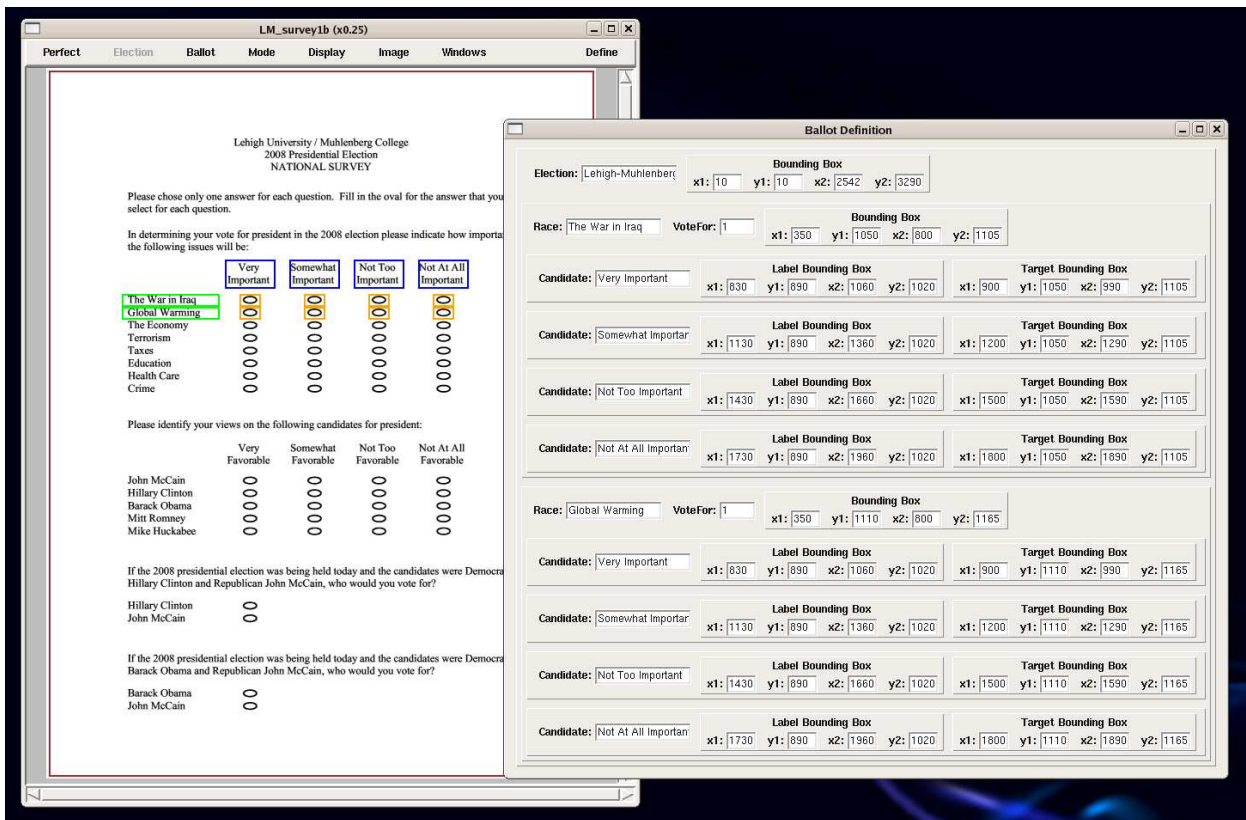


Figure 1. BallotTool system displaying the Lehigh-Muhlenberg simulated survey form.

on where a name appears on the ballot sheet. (It is already well understood that ballot layout can have an impact on voter behavior.)

- *High cost of manual recounts.* Recounting all of the ballots in a large geographic area can be expensive, both in terms of time and money.
- *Human error and human bias in performing audits and recounts.* While human ballot readers are deserving of more trust than machines, at least as of today, they also bring with them personal biases which may intentionally or unintentionally alter the outcome of an election. Recounts also present opportunities for collusion.
- *Computer “hackers” attempting to manipulate the vote.* This fear is the driving force behind the push toward paper ballots, but it should be noted that the electronics of optical scan systems have been proven to be just as vulnerable as DRE systems [7, 10].
- *Traditional ballot-box stuffing.* While there is no such thing as a perfectly secure voting system, some approaches are safer than others, a mantra that should be always kept in mind. Low-tech approaches have

undoubtedly resulted in the theft of more elections throughout history than the current cyberthreats that now receive so much media attention.

- *The need to preserve anonymity.* Many solutions that come to mind for securing and processing paper ballots place the anonymity of voters at risk. It is for this reason, for example, that current approaches for providing a Voter Verified Paper Audit Trail (VVPAT), an alternative that has been proposed to paper ballots, cannot be certified for use in Pennsylvania and certain other states. Since existing VVPAT printers maintain their audit trails sequentially, it would be possible to recover a voter’s votes by noting the order in which people use the machine and later obtaining a copy of the paper tape. Likewise, schemes for pre-printing a unique ID on each ballot also fall under suspicion.
- *Voter error.* As noted previously, the range of individuals who vote in a country’s elections reflects a broad spectrum of educational levels and literacy skills. Some voting technologies are more likely to induce errors than others; simply blaming the voter in all such cases is not appropriate. The Florida butterfly ballot, cited at the beginning of our paper, is a perfect

```

<election ID="election001" Election="Lehigh-Muhlenberg Survey"
  bb_x1="10" bb_y1="10" bb_x2="2542" bb_y2="3290">
<race ID="race001" Race="The War in Iraq" VoteFor="1"
  bb_x1="350" bb_y1="1050" bb_x2="800" bb_y2="1105">
<candidate ID="cand001" Candidate="Very Important"
  bbl_x1="830" bbl_y1="890" bbl_x2="1060" bbl_y2="1020"
  bbt_x1="900" bbt_y1="1050" bbt_x2="990" bbt_y2="1105">
<candidate ID="cand002" Candidate="Somewhat Important"
  bbl_x1="1130" bbl_y1="890" bbl_x2="1360" bbl_y2="1020"
  bbt_x1="1200" bbt_y1="1050" bbt_x2="1290" bbt_y2="1105">
<candidate ID="cand003" Candidate="Not Too Important"
  bbl_x1="1430" bbl_y1="890" bbl_x2="1660" bbl_y2="1020"
  bbt_x1="1500" bbt_y1="1050" bbt_x2="1590" bbt_y2="1105">
<candidate ID="cand004" Candidate="Not At All Important"
  bbl_x1="1730" bbl_y1="890" bbl_x2="1960" bbl_y2="1020"
  bbt_x1="1800" bbt_y1="1050" bbt_x2="1890" bbt_y2="1105">
</race>
<race ID="race002" Race="Global Warming" VoteFor="1"
  bb_x1="350" bb_y1="1110" bb_x2="800" bb_y2="1165">
<candidate ID="cand001" Candidate="Very Important"
  bbl_x1="830" bbl_y1="890" bbl_x2="1060" bbl_y2="1020"
  bbt_x1="900" bbt_y1="1110" bbt_x2="990" bbt_y2="1165">
'''

```

**Figure 2. Portion of the definition file for the Lehigh-Muhlenberg simulated survey form.**

example of this.

- *Interpretation of marginal markings.* The crux for much of what we are studying is that two different ballot readers – humans and/or machines – may interpret the same marking differently. Such markings are called “marginal,” which is, of course, a relative term. Whether or not the ballot includes explicit instructions for how it should be marked, and whether or not the voter follows such instructions, legislation is usually written in terms of *voter intent*. In other words, markings that appear to reflect a voter’s desires should not be disqualified for purely technical reasons.
- *Testing and certification of electronic voting systems.* While the federal and state governments ostensibly test and certify electronic voting systems before they can be used in real elections, such evaluations are rudimentary at best. In Pennsylvania, for example, optical scan systems are tested by running 12 ballots and confirming that the tallies are correct. The shortcomings of this current approach to government qualification was dramatically demonstrated when California and Ohio contracted with independent security consultants to test voting systems used in their states, only to find numerous serious security holes that had passed the original certification process [5, 20].

### 3. The BallotTool System

In this section, we describe the prototype BallotTool system we are building to support our research on paper-based ballot processing. BallotTool contains a collection of useful software components for manipulating ballot images and their associated metadata. The BallotTool graphical user interface (GUI) is written in the popular Tcl/Tk scripting language [21] with versions that run under both the Linux and Microsoft Windows operating systems, where it also makes use of the standard Netpbm open source toolkit for manipulating image files [19]. Although our planned work covers a broad range of questions, many of them are centered around the image of a single ballot. Rather than develop a separate program for each task, BallotTool provides common functionality that is shared across a number of different applications, as will be described in the remainder of this paper. See Figure 1 for a screen snapshot of BallotTool in one of its operating modes: defining a basic ballot layout.

Underlying the BallotTool system is an XML-like language we have developed for describing ballots and elections. This provides a common representational framework for all of the applications we plan to study. Input is assumed to be a blank ballot image in PDF or TIF format. Metadata is built up through human interaction with the system, or, in certain cases, generated automatically. Figure 2 shows

a fragment of the specification corresponding to the ballot from Figure 1.

### 3.1. Ballot Definition

In addition to specifying the bounding box coordinates for relevant regions on the page, ballot definition must describe the logical components (*i.e.*, the semantics) of the election in question. Briefly, an *election* consists of some number of *rac*es, and each race contains some number of *candidates*.<sup>1</sup> A voter might cast a vote for one or more candidates in each race. Some elections permit multiple votes in a given race, while other times this would be considered an “overvote” which invalidates all the voter’s choices in that race. Undervoting (casting fewer votes in a race than one is permitted) is also a possibility that must be accounted for, of course.

Our ballot definition file format is designed to be sufficiently expressive to handle all of our intended applications throughout the course of the project.

### 3.2. Ballot Interpretation

As noted earlier, a significant portion of our research surrounds the issue of voter intent and the ways it might be interpreted by human and machine ballot readers. The notion of “ground-truth” – that is, the single correct answer as determined by a human observer which the machine then tries to obtain – has less relevance here than it does in traditional document analysis experiments [16].

BallotTool supports the collection of vote/no-vote judgments from human judges in an intuitive, point-and-click fashion. As will be described shortly, the results from machine processing of ballot images can be displayed for human feedback as well. In concert with our ballot specification language, allowances are made for multiple conflicting interpretations for each mark.<sup>2</sup>

### 3.3. Synthetic Ballot Generation

Our research on paper ballot processing includes a number of data collection and generation activities that are described more fully in a companion paper [18]. Here we briefly note that while it is often possible to obtain blank ballot images, real ballots that have been marked by voters for use in an election are normally subject to legal restrictions. Barriers to obtaining “live” data are not unknown

<sup>1</sup>It should be understood that these terms are used abstractly. Candidates, for example, need not be human, rather, they are choices a voter makes in response to the question posed by a race. This point is illustrated by the example in Figures 1 and 2.

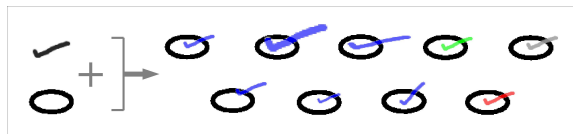
<sup>2</sup>Of course, in real elections, such conflicts must eventually be resolved through processes dictated by relevant legislation and/or the legal system.

in the field of document analysis research (scanned check images are likewise considered sensitive, for obvious reasons), but this may be one application area where the hurdles are insurmountable. Hence, we are currently conducting our own data collection activities using “simulated” ballots (*e.g.*, the Lehigh-Muhlenberg survey form shown earlier), as well as developing tools for generating large quantities of realistic-looking synthetic ballot images.

For generating synthetic ballot images, the relevant fields from the ballot specification include the race definitions (the candidates in each race, the maximum number of choices permitted, the rates at which each candidate receives a vote if the ballot is being generated randomly, etc.), as well as the physical locations of the appropriate mark targets (*e.g.*, the ovals which the voter is expected to fill in, indicated in orange in Figure 1).

In addition to the ballot substrate, ballot synthesis also requires a supply of previously drawn markings. This can be assembled in any of several ways: scanned off of paper ballots, drawn on-screen using a digitizing tablet, etc. Marks can be added to the ballot image either by superimposing them on the target (in this case, the mark must be drawn with a transparent background) or by replacing the target/mark combination as a single component.

Our software converts ballot and mark images to PNM format before they are manipulated. Image-based transformations are provided so that a single mark can be given a variety of visible manifestations. Marks can be scaled independently in the x- and y-dimensions, and they can be rotated by an arbitrary amount. The color of the mark can also be remapped before it is placed in the ballot image. Finally, the new ballot image is converted to TIF format using LZW compression. A PDF version (for printing purposes) is also created at the same time. Examples of the variations created from a single input checkmark are shown in Figure 3.



**Figure 3. Variations synthesized from a single input checkmark.**

Complete collections of synthetic ballots (*i.e.*, “elections”) possessing certain desired properties can, of course, be generated just as easily. The specification in this case includes the number of ballots (voters), selection rates for each candidate (including possible correlations between related candidates), overvote and undervote rates, etc.

## 4. Unbiased Visual Auditing

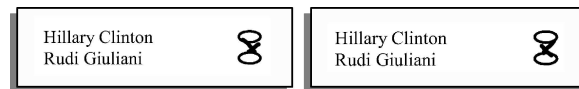
One of the goals we are pursuing in our research is the development of semi-automated tools to help eliminate human bias in auditing election results. Bias arises because people naturally have an inherent preference for particular candidates or issues.

Current procedures for manually recounting paper ballots can be regarded as largely social processes, with two or more representatives from competing political parties examining each ballot one-by-one and debating, sometimes heatedly, whether a given mark should be counted as a vote. While such a process is ostensibly intended to minimize the effects of bias, it clearly is suspect. Election fraud could be carried out by attempting to manipulate the angle at which a ballot is viewed by the other party, or by surreptitiously adding spurious marks to ballots, thereby invalidating them (because of conflicting marks) in subsequent recounts. If votes are counted on displays of digital ballot images instead, these subterfuges are rendered essentially unworkable. There remains, however, the possibility of voluntary or involuntary bias in counting.

Using our software for creating synthetic ballot images, two colleagues from the social sciences are now performing a study of how human evaluators judge ballot marks. The results demonstrate important variation in how people judge voter intent, an important criterion used in many state election laws. In some cases nearly identical marks are judged very differently by human evaluators. Figure 4 depicts a specific instance from this study, a small portion of a simulated survey form that purported to reflect a respondent's attitudes on a number of politically related issues. This establishes a context under which a test subject's evaluation of a marginal ballot marking may be biased. Establishing the determinants of such judgments is an important element of our ongoing research. In particular, we are examining the relative effects of the social characteristics of the judges (partisanship, age, etc.) and contextual information (votes on the same ballot on related questions) provided by different ballot designs in judgments human evaluators make regarding voter intent.

We are planning to conduct experiments that may lead to simple ways of avoiding such bias. These experiments will make use of the ballot images generated using the software described earlier.

One set of experiments will consist of comparing human vote counts using context versus human counts without using context (oddly called "blind recording"). Context is present when enough of the ballot is visible to the human counter to identify the party, candidate, or proposition for which a vote was cast. Context is absent when only the mark area is displayed. The screen snapshots in Figure 5 illustrates the steps required to eliminate context and enable



**Figure 4. Fragments from our study examining human interpretation of a simulated political survey. Left: interpreted as a vote for Clinton 15% of the time and a vote for Giuliani 2% of the time. Right: interpreted as a vote for Clinton 10% of the time and a vote for Giuliani 12% of the time.**

blind audits.

This figure again depicts the BallotTool software, this time running in a different mode. The topmost image (a) in Figure 5 shows the original ballot with all labels visible. The middle image (b) shows the ballot with the candidate labels obscured (recall that in an election, the candidates need not be human – they merely refer to the options presented to the voter). The lower image (c) is the one the human auditor actually sees. Here the voter choices have been randomly shuffled so that the human interpretation is based entirely on the visible marks (including other marks made by the voter on the same page), without any information about the actual choices the voter has made.

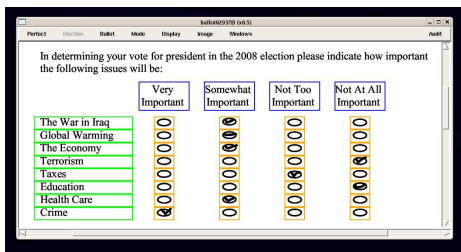
Also shown in the lower screen snapshot in Figure 5 is an enlarged closeup view of the mark that is currently under the cursor. This is provided as a aid to the human judge in interpreting the mark. The judge can indicate a vote/no-vote decision by clicking on the mark in question. As seen in the figure, several marks have already been annotated as votes; these appear with red outlines.

Semi-automating the processing of election recounts could address one of the most pressing and well-publicized issues in moving to paper support for electronic voting. While RES, the Remote Encoding System (human encoding of postal address images), has been an economic success for the U.S. Postal Service, we must take into account the additional legal and social constraints imposed in the case of elections. Whether an approach like the one we are proposing would be accepted as a replacement for the current (flawed) social process is a question we plan to explore.

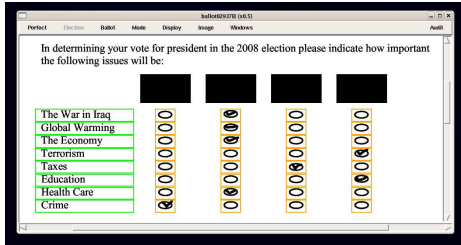
## 5. Homogeneous Class Display

Manually processing large quantities of ballot images in an audit or recount is a tedious and error-prone activity. Another area of our investigation is to explore approaches to facilitate detecting errors in machine or human interpretation of ballot images.

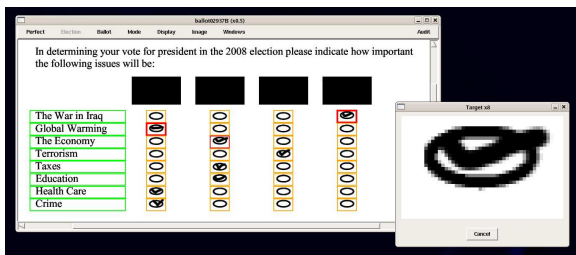
Our idea designed to verify automated mark sense counts is inspired by long-established methods of OCR verifica-



(a) Original ballot image.



(b) Candidate labels obscured.



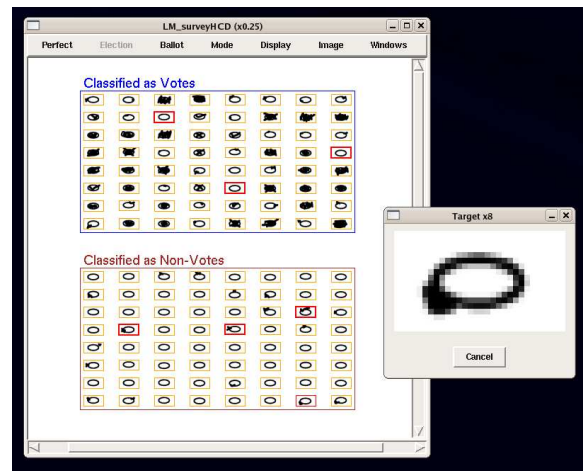
(c) Voter's selections randomly shuffled within races.

**Figure 5. Unbiased visual auditing.**

tion, homogeneous class display (HCD). Here the isolated images where a mark was registered as a vote are grouped for display. The images where a mark was not registered are similarly grouped. The positive marks are displayed in groups of, say,  $64 (= 8 \times 8)$  simultaneously, and the human operator is asked to note any discrepancies. The negative instances are displayed in similar groupings, and again the operator will notate any perceived discrepancies. This is depicted in the BallotTool screen snapshot Figure 6, where several potential recognition errors are evident on quick examination.

In a real application, spurious marks can also be introduced to assess operator error rate. (In some OCR service bureaus, the operators are notified of the presence of artificial errors in order to keep them alert.) HCD is, in principle, a very fast and convenient method of verifying automated counts, but it cannot be accepted without experimental evidence of its accuracy.

Following the practice of the postal RES, we intend to conduct variations on HCD that include displaying only targets that were not classified as marked or blank with very



**Figure 6. Homogeneous class display.**

high confidence. Experience with high-accuracy financial OCR and MICR (magnetic ink character recognition) indicates that reject/error ratios of the order of 100:1 or even 1,000:1 are necessary to achieve acceptable levels of undetected errors at the election precinct level. Since errors within a ballot are correlated, a more sophisticated reject strategy can be based on the confidence level of multiple marks on the same ballot.

Using synthetically generated (albeit artificially degraded) ballots avoids tedious ground-truthing. But we can still compare human accuracy on the paper ballots themselves with human accuracy on their images as a function of ballot quality, scan resolution, inspection speed, display size, and display quality. The resulting information should be extremely valuable for election planning and commercial development of improved voting machinery.

Auditing elections through electronic images of the paper trail offers certain benefits over auditing the paper trail itself. The main advantage is that image-based auditing can be done by physically separated election inspectors in parallel. Furthermore, it is also infinitely repeatable because the electronic image is preserved intact. While it might be argued that only the paper ballot itself is worthy of trust given the recent attacks demonstrated on electronic voting systems, this post-processing step could be developed independently of the proprietary positions taken by commercial vendors and hence is amenable to an open source model that would permit public scrutiny of the required software and hardware. Whether such auditing would prove acceptable to voters is a key question that we plan to examine on the social sciences side of our project.

These experiments are intended to demonstrate that the proposed procedures can guarantee the absence of systematic bias, and to determine how many duplicate inspections are required to reduce the expected number of random er-

rors to any pre-specified level that is acceptable.

## 6. Conclusions

In this paper, we have presented our BallotTool system which is designed to support a multi-institution, cross-disciplinary research project examining issues that arise when paper ballots are used in elections. While sharing some common traits with other document analysis applications, ballot processing is in many ways unique. We discussed some of the special constraints raised in reading ballots as opposed to more general document images, and described our approaches to several important problems, including eliminating human bias in the auditing of ballots and facilitating the detection of machine recognition errors. Current results can be found on the PERFECT website [22].

## 7. Acknowledgments

This work was supported in part by the National Science Foundation under award numbers NSF-0716368, NSF-0716393, NSF-0716647, and NSF-0716543. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect those of the National Science Foundation.

We wish to thank collaborators who are working with us on other aspects of this project, including Anne Miller (Rensselaer Polytechnic Institute), Ziad Munson and Nicholas Hinnerschitz (Lehigh University), and Christopher Borick (Muhlenberg College).

## References

- [1] Parties outline recount proposals with House control at stake, December 12 2006. [www.whptv.com](http://www.whptv.com).
- [2] SAT process review: Answer sheet processing (final report). Booz Allen Hamilton, May 26 2006. <http://www.collegeboard.com/sat/answersheetprocessing/report>.
- [3] The Machinery of Democracy: Protecting Elections in an Electronic World. Technical report, The Brennan Center Task Force on Voting System Security, June 27 2006. [http://brennan.3cdn.net/52dbde32526fdc06db\\_4sm6b3kip.pdf](http://brennan.3cdn.net/52dbde32526fdc06db_4sm6b3kip.pdf).
- [4] 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. <http://vote.nist.gov/DraftWhitePaperOnSlinVVSG2007-20061120.pdf>.
- [5] Secretary of State Top-to-Bottom Review of California Voting Systems, March 2008. [http://www.sos.ca.gov/elections/elections\\_vsr.htm](http://www.sos.ca.gov/elections/elections_vsr.htm).
- [6] A. J. Feldman, J. A. Halderman, and E. W. Felten. Security analysis of the Diebold AccuVote-TS voting machine. Technical report, Princeton Center for Information Technology, September 13 2006.
- [7] H. Hursti. The Black Box Report: Critical security issues with Diebold optical scan design. Technical report, Black Box Voting Project, July 4 2005.
- [8] H. Hursti. Diebold TSx evaluation: Critical security issues with Diebold TSx. Technical report, Black Box Voting Project, May 11 2006.
- [9] D. W. Jones. Counting mark-sense ballots, February 2002. <http://www.cs.uiowa.edu/jones/voting/optical/>.
- [10] A. Kiayias, L. Michel, A. Russell, and A. A. Shvartsman. Security assessment of the diebold optical scan voting terminal. Technical report, UConn Voting Technology Research Center, October 30 2006.
- [11] B. Klein, S. Agne, and A. D. Bagdanov. Understanding document analysis and understanding (through modeling). In *Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR'03)*, pages 1218–1222, Edinburgh, Scotland, August 2003.
- [12] B. Klein, S. Agne, and A. Dengel. Results of a study on invoice-reading systems in Germany. In S. Marinai and A. Dengel, editors, *Document Analysis Systems VI*, volume 3163 of *Lecture Notes in Computer Science*, pages 451–462. Springer-Verlag, Berlin, Germany, 2004.
- [13] B. Klein and A. R. Dengel. Problem-adaptable document analysis and understanding for high-volume applications. *International Journal on Document Analysis and Recognition*, 6(3):167–180, March 2004.
- [14] T. Kohno, A. Stubblefield, A. D. Rubin, and D. S. Wallach. Analysis of an electronic voting system. In *Proceedings of the IEEE Symposium on Security and Privacy*, May 2004.
- [15] M. Lausen. *Design for Democracy: Ballot + Election Design*. University of Chicago Press, 2007.
- [16] D. Lopresti and G. Nagy. Issues in ground-truthing graphic documents. In *Proceedings of the Fourth IAPR International Workshop on Graphics Recognition*, pages 59–72, Kingston, Ontario, Canada, September 2001.
- [17] D. Lopresti, A. Tomkins, J. Zhou, and J. Zhou. Systematic bias in OCR experiments. In *Proceedings of Document Recognition II (IS&T/SPIE Electronic Imaging)*, pages 196–204, San Jose, CA, February 1995.
- [18] G. Nagy, D. Lopresti, A. Miller, and E. B. Smith. Election ballot test data. Submitted for publication, 2008.
- [19] Homepage for the Netpbm toolkit, 2008. <http://netpbm.sourceforge.net>.
- [20] Ohio Evaluation & Validation of Election-Related Equipment, Standards & Testing (EVEREST), March 2008. <http://www.sos.state.oh.us/sos/info/everest.aspx>.
- [21] J. Ousterhout. *Tcl and the Tk Toolkit*. Addison-Wesley, 1994.
- [22] Paper and Electronic Records for Elections: Cultivating Trust (PERFECT), 2008. <http://perfect.cse.lehigh.edu/>.
- [23] M. A. Wertheimer. Trusted agent report: Diebold AccuVote-TS voting system. Technical report, RABA Innovative Solution Cell, January 20 2004.
- [24] K. Zetter. NH recount uncovers discrepancies – result of human error officials say. *Wired Blog Network*, January 22 2008. <http://blog.wired.com/27bstroke6/2008/01/nh-recount-unco.html>.
- [25] J. Zhou and D. Lopresti. Improving classifier performance through repeated sampling. *Pattern Recognition*, 30(10):1637–1650, 1997.