

Document Analysis Issues in Reading Optical Scan Ballots

Daniel Lopresti*
lopresti@cse.lehigh.edu

George Nagy†
nagy@ecse.rpi.edu

Elisa Barney Smith‡
EBarneySmith@boisestate.edu

ABSTRACT

Optical scan voting is considered by many to be the most trustworthy option for conducting elections because it provides an independently verifiable record of each voter's intent. While op-scan technology has been in use for decades, attempts to improve the machine-reading of ballots raise a range of interesting issues in document image analysis. Work thus far has been hindered by a lack of real-world data, however, since ballots associated with actual elections are kept secure from the public and normally destroyed after a period of time. Fortunately, as a result of a recent challenged federal election in Minnesota, a large number of op-scan ballot images were made available for public inspection on the Web.

In this paper, we present the Minnesota op-scan ballot collection as a unique resource to the document analysis community. We discuss important considerations regarding the definitions of a legal vote and a valid ballot which cannot be ignored for the purposes of technical expediency. Our efforts to annotate the collection are also described, including the development of a graphical tool for collecting ground-truth interpretations and the protocol now being employed. The collection, consisting of ballot images, file formats, and associated truth data for part of the set, is being made openly available to facilitate research in this important area.

Categories and Subject Descriptors

I.7.5 [Document and Text Processing]: Document Capture—*document analysis*

General Terms

human factors, experimentation, measurement, reliability

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

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

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

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

DAS '10, June 9-11, 2010, Boston, MA, USA

Copyright 2010 ACM 978-1-60558-773-8/10/06 ...\$10.00

1. INTRODUCTION

Spurred by problems that occurred during the contested 2000 U.S. Presidential Election, the process by which voting is conducted has undergone major changes over the past ten years. Soon after the resulting push to adopt electronic voting equipment, computer security experts and concerned citizens began raising serious questions about the reliability and trustworthiness of such systems [4, 6]. Direct Record Electronic (DRE) voting, once seen as an obvious solution to the problems that occurred in Florida in 2000, is now viewed by many as an unacceptable compromise [5, 16]. The tide is now turning toward 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 extensively studied in the past [8, 10]. Mark-sense readers preceded OCR by several decades: the IBM 805 Test Scoring Machine was introduced in 1937. The bubbles were relatively large and the marks had to fill most of the bubble in order to be reliably 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).

Despite its advantages, introducing (or, rather, re-introducing) paper into the modern election process raises other issues that merit the attention of the document analysis research community. In Chester County, Pennsylvania, for example, 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]. The same recommendation arose in Booz Allen Hamilton's analysis of a well-publicized problem in scoring the October 2005 Standardized Achievement Test, taken yearly by millions of college-bound high school students, which caused a number of students to receive scores much lower than they deserved [3]. 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 [23]. Finally, in perhaps the most well publicized case

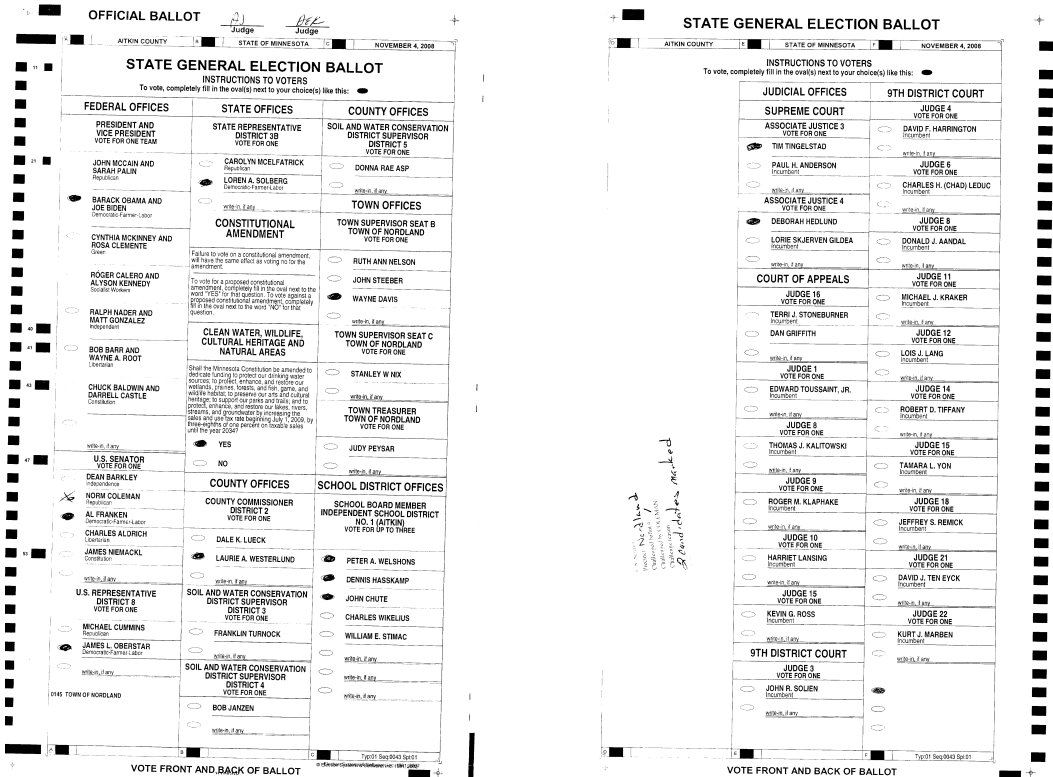


Figure 1: Front and back of one ballot from the collection.

of recent note, the extremely close 2008 U.S. Senate race in Minnesota ended in a challenge between two candidates, Republican Norm Coleman and Democrat Al Franken, that resulted in a public recount where over 6,000 op-scan ballot images were posted on the World Wide Web for the inspection of anyone who was interested [14].¹

Processing paper ballots used in elections differs from other document analysis tasks in important ways. The range of individuals who use a paper ballot is likely to be much greater than in typical 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, legally mandated, requirement is that ballots must preserve a voter's anonymity. This precludes including unique identifiers (e.g., a serial number) 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. The officials who administer elections are volunteers with little or no specialized training in operating the equipment in question. 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,

¹Franken was ultimately declared the winner after months of legal wrangling. The final margin of victory was determined to be 312 votes out of 2.9 million cast.

there is tremendous public interest; the process of casting and counting votes must be transparent and trustworthy.

Research on op-scan ballot reading has been hindered to date because of a lack of access to real data. While it is possible to create synthetic ballot images, and we have done so ourselves [18], concerns must be raised about the validity of such an approach because many of the problems encountered in elections involve unanticipated behavior on the part of voters. The large-scale release of ballots from the 2008 Minnesota Senate race has addressed this issue in a resounding fashion; in this paper we describe our efforts to collect, organize, and annotate the 6,737 op-scan ballots that were made available to the public. We also provide an overview of the system we have developed to support the efficient ground-truthing of ballot images, an early prototype of which was the subject of our paper at the 2008 DAS Workshop [12].

Before beginning research in a new problem area, it is important to know the "ground-rules." Certain kinds of simplifying assumptions employed in other applications (e.g., forms reading) are inappropriate when it comes to processing scanned ballots. Voters cannot be disenfranchised just because they fail to follow instructions; if the voter's intent is clear to a human judge, then the machine must try to match this. Hence, another contribution of this paper is highlighting challenging cases from real ballots that a proposed solution ought to be able to handle. While we have begun to develop our own specific methods for certain sub-problems, these appear elsewhere (e.g., [2, 21]).

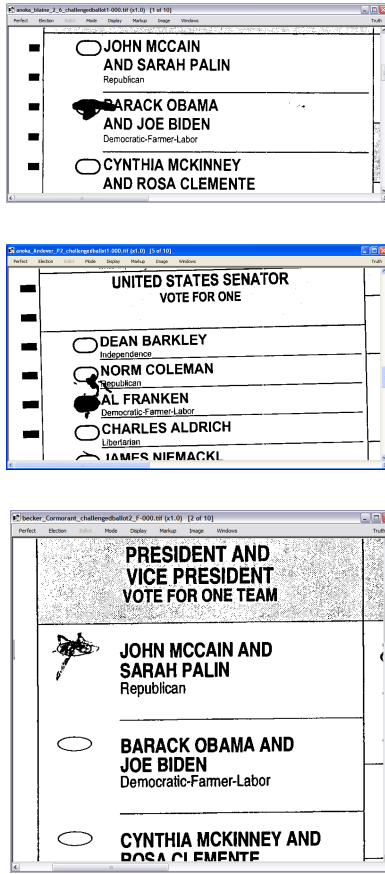


Figure 2: Sloppy-but-valid marks #1.

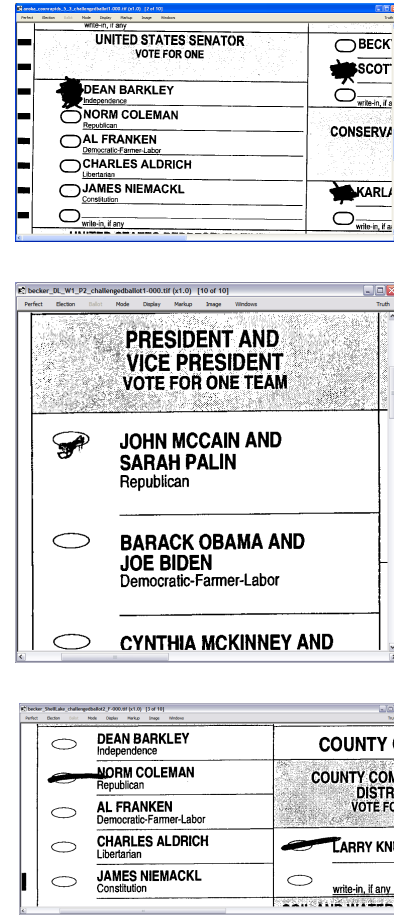


Figure 3: Sloppy-but-valid marks #2.

The remainder of the paper presents work in progress and is organized as follows. Section 2 describes the dataset of real-world ballots we are assembling for release to the document analysis research community. We highlight a number of interesting cases we have encountered in perusing the collection. In Section 3, we discuss the tool we have developed for ground-truthing ballot images. Our file formats and truthing protocol are described in Section 4. Finally, Section 5 concludes with a discussion of ongoing work.

2. DATASET

To our knowledge, no large-scale dataset of voter-marked ballots from a real election has ever been made openly available to a research community until now. In this section, we describe the ballot collection and the circumstances under which it came to be released to the public. As per the goals of the special focus on contributed datasets at the DAS 2010 workshop, the complete set of ballot images along with ground-truth and other associated metadata will be submitted for dissemination via the IAPR TC-11 website [7].

2.1 The 2008 Minnesota Senate Race

The 2008 General Election in the United States took place on November 4. Among the races decided that day was the presidential election in which Barack Obama was elected the 44th U.S. President. In the State of Minnesota, in addition to the presidency and a number of state-wide and local

races, citizens also voted to elect a U.S. Senator. Five candidates were listed on the ballot. In the initial tally, Republican Norm Coleman received 1,211,590 votes (41.988% of the votes cast) while Democrat Al Franken received 1,211,375 votes (41.981% of the votes cast). Because of the closeness of the race, a mandatory recount was ordered [19]. After a series of challenges, Franken was declared the winner [20].

In the process of performing recounts, representatives from either candidate may *challenge* a given ballot for not meeting the legal requirements set by the state. The intention is to deny the opposing candidate a vote. Under Minnesota law, a vote must be counted if it is possible to determine intent, even if the voter failed to follow instructions. So while the directions on the ballot read quite simply “To vote, completely fill in the oval(s) next to your choice(s),” a vote should still be counted if the voter uses, say, a check mark to indicate his/her choice. Ballots are not counted for a particular race if more candidates are marked than allowed in that race (with no attempt to strike any of them out). The entire ballot is invalidated if there is any mark on it that would identify the voter.

2.2 Scanned Ballot Images

In the process of performing the recount, ballots that had been challenged were scanned and placed online so that

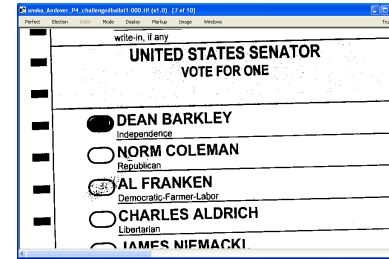
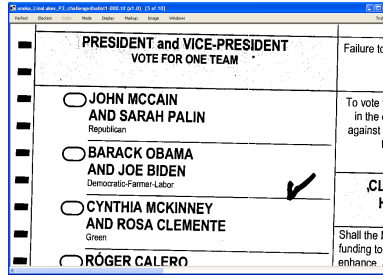
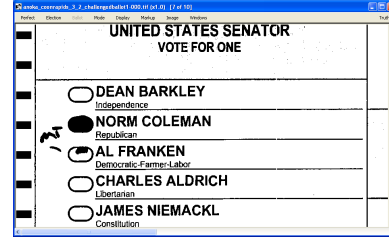
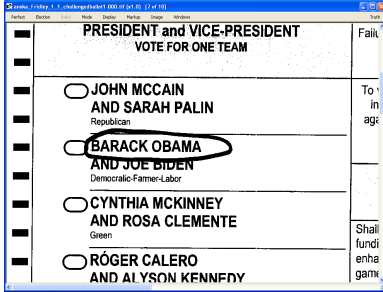
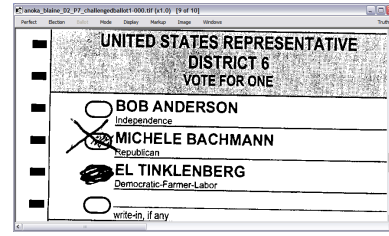
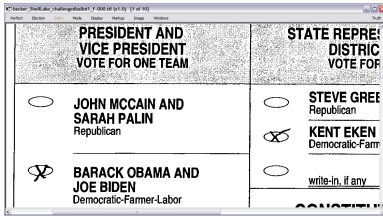


Figure 5: Attempts to cancel a vote.

Figure 4: Non-conforming marking styles.

the public would have an opportunity to view them. A number of websites made the ballot images available, including the site for Minnesota Public Radio [14]. While a short video demonstrating the scanning process can be found on YouTube [22], only minimal technical details can be deduced. Ballots were first photocopied and the originals stored in a secure location. The photocopies were then scanned to PDF using an auto-feeder equipped flatbed scanner. The ballot was two-sided, with both sides scanned simultaneously.

To collect all of the ballots from the MPR website, we wrote a simple web “crawler” that automatically downloaded the files, saving them under their original file names. Another program was then used to extract the images from the PDF, saving the front and the back of each ballot as a separate TIF file. There are a total of 6,737 ballots in the set. Examination of the TIF suggests that the ballots were scanned at 300 dpi bitonal, and that lossy compression was never used in the handling of the files. Hence, they form an ideal dataset for document analysis research.

Figure 1 shows the front and back of a challenged ballot from Norland Township in Aitkin County, MN. This particular ballot was challenged by representatives for Coleman who claimed that the voter marked two candidates in the Senate race. In this case, it appears as though the voter intended to vote for Franken, but mistakenly started to fill in the oval

for Coleman before deciding to cross it out.

Ideally, a system built for interpreting op-scan ballots should replicate the same understanding of voter intent possessed by a knowledgeable human judge. In examining the collection of ballot images and contemplating the associated pattern recognition problems, a range of issues become evident. We have, for example, encountered: targets that are incompletely filled-in or where the marking extends well beyond the boundary of the oval (Figures 2-3), non-conforming marking styles (Figure 4), attempts to correct a mistaken vote by crossing it out or erasing it (Figure 5), intended votes that look like cancelled votes (Figure 6), stray marks on the ballot that might be confused with votes (Figure 7), bleed-through from one side of the ballot to the other (Figure 8), and, finally, handwritten annotations and other markings that are considered to be “identifying” and hence which invalidate the ballot (Figure 11). Many of these examples arise in races other than the Senate contest.

It must be noted that this is most certainly not a random sampling of ballots: whether the challenge was deemed valid or not, each of these ballots was questioned for some reason. It is highly likely that the vast majority of the ballots cast in Minnesota in 2008 were well-marked and would create no difficulty for even a simple ballot-reading algorithm. On the other hand, as the senate race in that state demonstrated, even a small error rate ($< 0.01\%$) could alter the outcome of an important election with national implications.

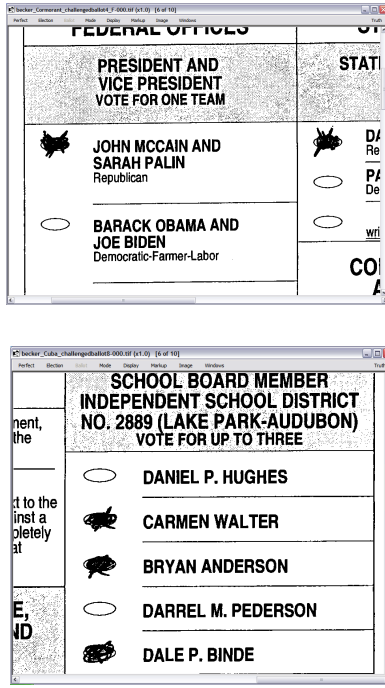


Figure 6: Intended votes that appear cancelled.

3. BALLOTTOOL SYSTEM

Ground-truth for op-scan ballots is different from the truth employed for, say, page layout analysis or character recognition. In addition to the markings on the page, relationships between markings (the semantics of the voter's intent) must also be captured. As described in our earlier paper [12], we have built a graphical tool to support the ground-truthing of ballot images. In this section, we describe the current BallotTool system as it is being used to annotate the Minnesota Challenged Ballots Dataset. 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 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. See Figure 9 for a screen snapshot of BallotTool displaying a partially annotated ballot image.

BallotTool is designed to support browsing and mark-up of sets of ballot images. It collects user judgments in an intuitive, point-and-click fashion. Annotation takes place at two different levels of abstraction. At the lowest level, the user can draw bounding boxes for an open-ended assortment of object types; currently this includes valid votes, cancelled votes, stray marks, hand stamps, and handwriting by the voter or officials. In addition, arbitrary text labels can be placed anywhere on the ballot image. As part of the ground-truthing process, users are also asked to mark the location of fiducials when they appear on the ballot, as well as to draw a line over a ruling on the image they know to be horizontal as a mechanism for estimating page skew.

The higher level of abstraction allows users to associate

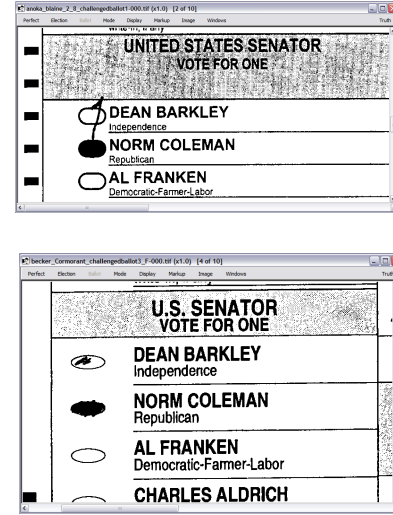


Figure 7: Stray marks on the ballot face.

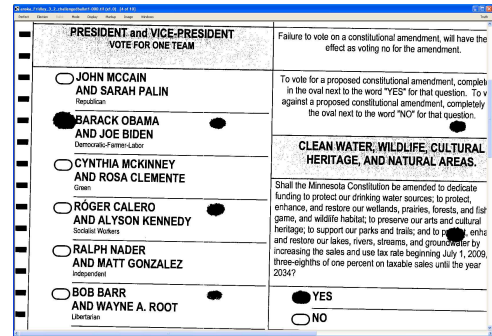


Figure 8: Bleed-through from ballot back to front.

mark-up on the ballot with the candidate it belongs to, and to indicate which candidate(s), if any, receive the legal vote(s) in a given race. Both levels of abstraction are necessary to characterize fully each ballot in the collection.

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 “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. In concert with our ballot specification language, allowances are made for multiple conflicting interpretations for each mark. All user interactions with BallotTool are logged for later analysis.

4. GROUND-TRUTH FORMATS

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, including the current dataset. Meta-data is built up through human interaction with the system, or, in certain cases, generated automatically. Figure 10 shows a fragment of the specification corresponding to the ballot from Figure 9.

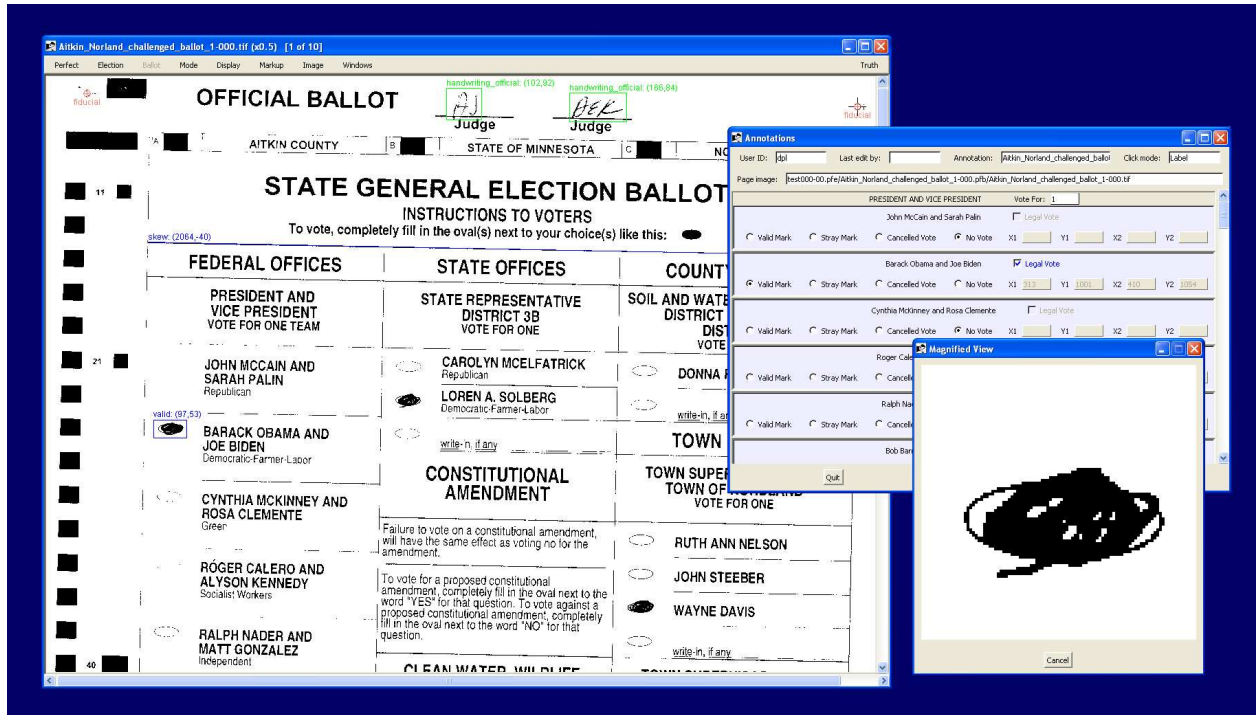


Figure 9: BallotTool system displaying a partially annotated ballot from Aitkin County.

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*.² 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. Under-voting (casting fewer votes in a race than one is permitted) is also a possibility that must be accounted for, of course.

Lower-level markup (*e.g.*, the location of stray marks) is maintained as a separate collection of objects stored in a simple geometric representation. The ballot file formats are designed to be sufficiently expressive to handle all of our intended applications throughout the course of the project.

Our protocol for ground-truthing is described in a separate 15-page document [11]. This includes instructions for installing the BallotTool software and its supporting utilities, the procedure to be followed for truthing, rules for interpreting markings, and illustrative examples. The standard for voter intent is set by the Minnesota Statutes [15].

To date, eight subjects have annotated a total of 780 ballots, with some intentional overlap to allow us to examine the consistency between users. This data collection is ongoing and we plan to present additional details regarding the ground-truthing activity at the DAS workshop.

ing and we plan to present additional details regarding the ground-truthing activity at the DAS workshop.

5. DISCUSSION

The task of assembling and ground-truthing a large collection of real optical scan ballots is instructive in its own right. While certain aspects of this work are specific to one particular election that took place in the United States in 2008, many of the questions we are studying have wider implications; assumptions and methods proven unreliable in one locale cannot be trusted in any other. In this paper, we surveyed some of the motivating factors behind our project, discussed the special constraints raised in processing ballots as opposed to more general document images, and described the current status of our BallotTool system. Further details can be found on the PERFECT project website [17].

More broadly, we take this opportunity to briefly highlight some of the basic problems we are working to address:

- *Undetected failures in the machine-reading of ballots.* There is usually no warning when recognition errors arise in optical scan systems [1, 3]; processing the ballot a second time may lead to a different result [24].
- *Systematic errors due to ballot layout.* Our past work in OCR demonstrated that recognition errors are not uniformly distributed across the page [13]; the same observation may be true of ballots, a fact which may disadvantage one candidate over another based purely on where a name appears on the ballot sheet.
- *High cost of manual recounts.* Recounting all of the ballots in a large geographic area can be expensive,

²It 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. The Minnesota ballot shown in Figures 9 contains a Constitutional Amendment where the two candidates are “yes” and “no.”

```

<annotations ver="2.0">
<race000 race="PRESIDENT AND VICE PRESIDENT" votefor="1" columns="8" background="#eeeeff">
<cand000 name="John McCain and Sarah Palin" variable="legalvote" value="0">
<radiobutton text="Valid Mark" variable="John_McCain_and_Sarah_Palin" assign="valid" value="novote">
<radiobutton text="Stray Mark" variable="John_McCain_and_Sarah_Palin" assign="stray" value="novote">
<radiobutton text="Cancelled Vote" variable="John_McCain_and_Sarah_Palin" assign="cancelled" value="novote">
<radiobutton text="No Vote" variable="John_McCain_and_Sarah_Palin" assign="novote" value="novote">
<entry text="X1" textvariable="x1" value="">
<entry text="Y1" textvariable="y1" value="">
<entry text="X2" textvariable="x2" value="">
<entry text="Y2" textvariable="y2" value="">
</cand000>
<cand001 name="Barack Obama and Joe Biden" variable="legalvote" value="1">
<radiobutton text="Valid Mark" variable="Barack_Obama_and_Joe_Biden" assign="valid" value="valid">
<radiobutton text="Stray Mark" variable="Barack_Obama_and_Joe_Biden" assign="stray" value="valid">
<radiobutton text="Cancelled Vote" variable="Barack_Obama_and_Joe_Biden" assign="cancelled" value="valid">
<radiobutton text="No Vote" variable="Barack_Obama_and_Joe_Biden" assign="novote" value="valid">
<entry text="X1" textvariable="x1" value="313">
<entry text="Y1" textvariable="y1" value="1001">
<entry text="X2" textvariable="x2" value="410">
<entry text="Y2" textvariable="y2" value="1054">
</cand001>
...

```

Figure 10: Portion of the ground-truth file for the ballot shown in Figure 9.

both in terms of time and money.

- *Human error and human bias in performing audits and recounts.* While human ballot readers deserve 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.
- *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 [6, 9].
- *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 certain states. 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.

- *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, 16].

6. ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation under awards 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. Daniel Lopresti also acknowledges support from a DARPA IPTO grant administered by Raytheon BBN Technologies, and George Nagy acknowledges NSF support under award NSF-044114854. Lastly, we thank the reviewers for their helpful feedback.

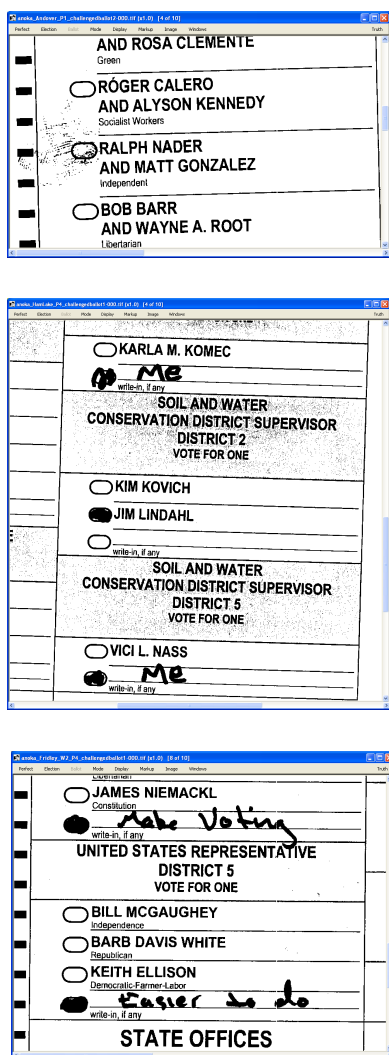


Figure 11: Invalid ballot due to identifying marks.

7. REFERENCES

- [1] Parties outline recount proposals with House control at stake, December 12 2006. <http://www.whptv.com>.
- [2] E. Barney Smith, D. Lopresti, and G. Nagy. Ballot mark detection. In *Proceedings of the Nineteenth International Conference on Pattern Recognition*, pages 4 (CD-ROM), Tampa, FL, December 2008.
- [3] SAT process review: Answer sheet processing (final report). Booz Allen Hamilton, May 26 2006. <http://www.collegeboard.com/sat/answersheetprocessing/report>.
- [4] The machinery of democracy: Protecting elections in an electronic world. The Brennan Center Task Force on Voting System Security, June 27 2006. <http://brennan.3cdn.net/52dbde32526fdc06db4sm6b3kip.pdf>.
- [5] Secretary of State Top-to-Bottom Review of California Voting Systems, March 2008. http://www.sos.ca.gov/elections/elections_vsr.htm.
- [6] H. Hursti. The Black Box Report: Critical security issues with Diebold optical scan design. Technical report, Black Box Voting Project, July 4 2005.
- [7] IAPR TC-11 Website, April 2010. <http://www.iapr-tc11.org>.
- [8] D. W. Jones. Counting mark-sense ballots, 2002. <http://www.cs.uiowa.edu/~jones/voting/optical/>.
- [9] 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.
- [10] M. Lausen. *Design for Democracy: Ballot + Election Design*. University of Chicago Press, 2007.
- [11] D. Lopresti. Instructions for ground-truthing opscan ballot images, 2010. http://perfect.cse.lehigh.edu/Documents/GTInstructions_1-18-10.pdf.
- [12] D. Lopresti, G. Nagy, and E. Barney Smith. A document analysis system for supporting electronic voting research. In *Proceedings of the Eighth IAPR Workshop on Document Analysis Systems*, pages 167–174, Nara, Japan, September 2008.
- [13] D. Lopresti, A. Tomkins, J. Zhou, and J. Zhou. Systematic bias in OCR experiments. In *Proceedings of Document Recognition II*, pages 196–204, San Jose, CA, February 1995.
- [14] Minnesota Senate Recount: Challenged ballots: You be the judge, 2008. http://minnesota.publicradio.org/features/2008/11/19_challenged_ballots/.
- [15] 2009 Minnesota Statutes: 204C.22 DETERMINING VOTER'S INTENT. <https://www.revisor.mn.gov/statutes/?id=204C.22>.
- [16] Ohio Evaluation & Validation of Election-Related Equipment, Standards & Testing, March 2008. <http://www.sos.state.oh.us/sos/info/everest.aspx>.
- [17] Paper and Electronic Records for Elections: Cultivating Trust (PERFECT), 2010. <http://perfect.cse.lehigh.edu/>.
- [18] PERFECT Project: Ballot Research, 2010. <http://perfect.cse.lehigh.edu/BallotResearch.html>.
- [19] T. Scheck and T. Nelson. FAQ on the Senate recount. MPRnewsQ, November 6 2008. http://minnesota.publicradio.org/display/web/2008/11/06/recount_faq/.
- [20] United States Senate election in Minnesota, 2008. Wikipedia, April 2010. http://en.wikipedia.org/wiki/United_States_Senate_election_in_Minnesota,_2008.
- [21] P. Xiu, D. Lopresti, H. Baird, G. Nagy, and E. Barney Smith. Style-based ballot mark recognition. In *Proceedings of the Tenth International Conference on Document Analysis and Recognition*, pages 216–220, Barcelona, Spain, July 2009.
- [22] Scanning Challenged Ballots In MN US Senate Recount (YouTube Video), April 2010. <http://www.youtube.com/watch?v=wnyky2A-Z7I>.
- [23] 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>.
- [24] J. Zhou and D. Lopresti. Improving classifier performance through repeated sampling. *Pattern Recognition*, 30(10):1637–1650, 1997.