

Automated Authentication of Current Identity Documents

Theodore Kuklinski, Ph.D.

Director of Research, AssureTec Systems, Inc., Manchester, NH

Abstract: *Much has been said about the difficulties in screening persons for possible identity fraud or security concerns based upon use of current driver's licenses or passports. The most often reasons given are the lack of standardization of security features and the layout for these documents. This criticism is focused on the inability of even a trained person to recognize valid documents and the specific parameters for each of these documents. In this paper, the focus is on the value of machine screening of the identity documents in circulation. The distinction is between human screening and the power of machine processing. The diversity of the identity documents and the issuer's attempts to exert their own unique identity for their documents is actually a benefit to machine screening. The rich variety of specific layout and production characteristics provide many examination points for evaluation. The processing power, storage capacity, and imaging options, only recently available at a reasonable price point, make real-time examination of all of the unique properties and a subsequent risk analysis of the results a practical approach.*

One of the most visible changes in society today is the need to present identity documents (ID's) in many more situations. A key objective in improving public security is the interdiction of individuals using counterfeit or stolen ID's to cross borders, use public transportation, open bank accounts, or enter facilities. However, security personnel are presented with an overwhelming number of identity documents and have only seconds to examine them, verify their authenticity, and approve the presenter.

New generation reader/authenticator/validator (RAV) technology can assist in the ID screening process for the wide variety of existing identity documents such as passports and drivers licenses. Such devices can read the information on the ID, authenticate the ID, and provide a security risk analysis. Acting as an inspector's automated assistant, their use permits an inspector to focus on evaluating the behavior of the presenter while the reader handles the detailed analysis of the document. A much more thorough ID inspection is possible, checking for many more security features, some of which are not accessible to the inspector without special equipment. More ID presenters can be processed faster with fewer inspectors at lower overall cost and with higher security confidence.

Almost 100,000 fraudulent documents were intercepted at U.S. ports of entry in 1998 [1]. How many are not being intercepted? The technology to create reasonable forgeries of current identity documents is both affordable and available. There are still many older style, unexpired, laminated licenses (e.g. New Jersey) in circulation; forging them requires little more than an inexpensive scanner, printer and laminator. Equipment capable of printing on the plastic card stock used for most ID cards today is now well within the budget of the average personal computer user. There are a variety of publications [2, 3] and websites that describe techniques for creating false identity documents. Other sources, rather than facilitating getting a counterfeit document, provide official looking secondary ID's that give the impression that they *could* be official documents, complete with genuine looking security features. They rely on the fact that document inspectors are not familiar enough with valid issued documents to recognize the bogus ones.

In the US, there have been calls for a national identity card with the idea that standardization of layout and security features will lead to better security; a standardized card would be easier to authenticate by the average inspector. Likewise, the perceived difficulty of validating current issue documents has spawned a movement toward greater use of biometrics in conjunction with ID's. While this is a worthy goal, it requires the step of enrollment to collect the biometric information. It will likely be a long time before currently issued documents are expired and replaced by "more secure" biometric documents. There has been a request [4] to push off the deadline for biometric passports from "visa waiver" countries until 2006.

In recent years, there has been a definite trend toward more secure identity documents, particularly driver's licenses. American states and Canadian provinces have largely converted over to more secure ID's incorporating such security features as ghost photos, check digits, security laminates, holograms, micro-printing, or patterns visible only in ultraviolet (UV), near-infrared (IR), or retro-reflective light and biometric features, magnetic stripes and barcodes. This trend also applies to passports and visas, which for many years have had a machine readable zone (MRZ) and a somewhat standardized layout for photos and other information.

The American Association of Motor Vehicle Administrators (AAMVA) organization provides standards for DL/ID (Driver's License/Identification) cards [5] and makes recommendations for placement of information on the card and the use of security features. Even within these guidelines, ID's issued by different states vary greatly in appearance. AAMVA admits "The increased use of the card for purposes other than proof of the privilege to drive have greatly increased the motivation to alter or counterfeit the DL/ID card." There is the desire, just as with state license plates, to impart some distinctive local identity. Sometimes license documents are not necessarily designed with readability in mind, for either the human inspector or machine readers, or in a manner that takes most advantage of the security features that may be available.

On the passport and visa front, the International Civil Aviation Organization (ICAO) has issued standards [6] for passports, visas, and other ID cards. For some 20 years, the ICAO has recommended the use of Machine Readable Zones (MRZ's), printed using the OCR-B font, in order to facilitate machine reading of such documents and as many as 300 million machine readable travel documents have been issued based on MRZ's. ICAO has established standards for the other areas of documents such as passports but there is still a large amount of discretion available to governments for customization of a unique appearance and the use of individual security features. Nonetheless, many passports and visas are non-standard or not machine readable (many even handprinted), and these will be circulating for years.

Screeners may become adept at recognizing the most common variants, but it is unreasonable to expect that security personnel or gate agents can memorize the detailed features found on the thousands of types of ID's presented for verification. Subtle design changes or even entirely new document designs are issued frequently. It is difficult for the inspector to keep up with these changes. Asa Hutchinson, Under Secretary for Border and Transportation Security at the Department of Homeland Security (DHS), in testimony before the U.S. Congress [7], acknowledges the problem: "there are more than 240 different types of valid driver's licenses issued within the United States" and further admits that "it would not be easy for CBP inspectors to have a passing familiarity with, let alone a working knowledge of, each of these documents. " Consider driver's licenses in the U.S. In addition to the standard issue driver's license and its older issue unexpired version, there are non driver ID's, Commercial, Provisional, Temporary, Under 21, Moped, Boat, and a host of other variants. Maryland, for example, has over 20 old and new license variants, with many of the variations printed in different colors.

So it remains that, for each document presented, an inspector must quickly know which visible security features to check and must instantly know where to look on a document to pull out necessary information such as the expiration date. There is no uniform date format for driver's licenses which is important in age verification situations. Most existing documents contain the same basic types of information. Unfortunately, the locations and format of such information varies widely with document type. The inspector must decode this information efficiently, reading small print in often poor and variable lighting conditions. They also need to be able to compare the photo on the ID and the face of the presenter or perhaps match the name on the ID and an airline ticket.

What aids are available to the person inspecting ID's to assist them in recognizing the wide variety of ID's that may be presented to them? Manuals [8] are available to law enforcement agencies and businesses, such as those serving alcohol, for the purpose of checking U.S. and Canadian driver's licenses. Issued annually, these so called "bar guides" commonly display examples of the current and unexpired past issue licenses and list some common

features to check. They may not contain many of the license variants that exist, particularly "Under 21" licenses, which are issued in a vertical format by many states.

While such guides may indicate that there should be a UV pattern, there is no indication of what that pattern is. How likely is the inspector to hold up a long line while they consult reference material to validate their fuzzy memory of some document feature? Equivalent services in the form of publications, software, and web reference sites, are available for passport and other international documents. Such resources draw a fine line between providing enough information for someone to validate a document, but not such complete information that a forger could produce a very good fake ID from the information provided.

We have seen the burdens that are put on document inspectors. Could a machine reader provide some assistance in this critical task? Machine readers have been available for some time that can read passports and other ID's that have MRZ's. Future identity documents will likely be equipped with many more features to make machine reading of them more efficient. We already are seeing the growth of smart card technology in ID's and magnetic stripes and barcode have been available for some time already. Nonetheless there are still a great number of "legacy" ID's that it would be useful and economically desirable to read.

Machine readers amplify the inspection ability of inspectors by providing automatic eyes on aspects of the presented ID they would be hard pressed to get otherwise. There may be a tremendous variability in the experience level of inspectors. They are subject to the many external factors – distraction, inattention, boredom, and even bribery. With the quality level of fraudulent ID's so high, the cursory glance of even experienced inspectors may easily fail to pick up the minor variations that could be telltale signs of document fraud. In most human inspections, typically an easily performed UV check is not performed. Machine readers don't get tired and can check all relevant details automatically and quickly. A machine reader can alert the human check to precisely those aspects of the document that may require closer inspection by providing a risk score. Their use can be an adjunct to the human inspection process, freeing them to focus on the behavior aspects of the presenter. In many checking situations, the focus is actually on the task of insuring that the face on the document matches that of the live person, given the vagaries of hairstyle, glasses, or facial jewelry.

A new generation of reader devices is available now, capable of fast full page color reading of passports and other identity documents. They feature automatic document sensing of up to passport sized documents (including driver's licenses), imaging in visible, IR, UV, and other lighting conditions, are trainable, and capable of scanning, reading, and authenticating in a few seconds. A modern reader system typically consists of a video camera, controllable lighting system, and a processing unit. The processing power, storage capacity, and imaging options, only recently available at a reasonable price point, make

real-time examination of all of the unique properties and a subsequent risk analysis of the results a practical approach. Document validation can be performed in a few seconds, an important factor where there may be long lines of people to be processed in a short time.

The diversity of identity documents and issuer's attempts to exert their own unique identity for their documents is actually a benefit to machine screening. The rich variety of specific layout and production characteristics provide many examination points for evaluation. An ID can easily be analyzed in a top-down fashion. Once the specific type of ID and particular issue are known, then one can look to a knowledge base of specific examination features that can be associated with that document. The position of particular fields may vary between issues. Certain unusual fonts may be used. Micro-printed areas are present in some modern licenses and can be highlighted. Under UV lighting, there is often a visible colored pattern or repeating pattern. In order to make it easier for machine reading of information, important text fields may be printed in ink which will be IR visible, allowing the scenic background information on many ID's to "drop out".

For speed of analysis, CCD color cameras can be used to capture the image. Real time video capture is feasible with the use of IEEE 1394 or USB-2 interface connections. This has the advantage of no moving parts, unlike scanners where either the document or reading head moves, requiring significantly more time for image capture. In the same time period, a camera based system can take several pictures of the document, each at different exposure settings and under different lighting conditions.

Older passport readers needed to capture only a portion of the document, typically just the MRZ. Now an entire passport page can be imaged in full color at sufficient resolution to enable accurate OCR of information fields, and even for reading barcodes. Cross checking of data derived from the MRZ and the data in the non MRZ region is possible. One of the driving forces toward full page reading is that it allows the automatic extraction of the photo which is important for matching against the live subject or checking against watchlists. With a full page document read, the image can be immediately displayed to the inspector. There need be very little time lost in the throughput process since the inspector could immediately inspect the color image of the document just as they might view the physical document. Smaller sized documents, such as driver's licenses, can be captured with the same system.

Older MRZ readers needed to capture only a binary image for OCR purposes. Earlier full page readers worked with gray scale images which didn't require as much processing power. However, color information is important for authenticating today's ID's. There is much information in a color image that can be used to identify the type and issue of an ID, and for analyzing various security features (e.g. multicolored UV patterns). Color filtering and image processing can be performed to enhance any of the information fields for OCR or other purposes.

Modern ID's usually contain security features that require the document to be illuminated under a number of different lighting conditions. Camera based readers make the task of utilizing multiple lighting sources in a short time feasible. They may use uniform white light for the capture of the visible image, near Infrared (B900) light for reading carbon based inks and security features without the color background, ultraviolet (UV) for detecting overlay patterns printed in UV sensitive inks, retro-reflective light for reading special laminates, and other specialized lighting for such features as holograms. With camera based systems, it is a matter of capturing an image frame under each of the lighting conditions, setting the exposure and gain, and switching the lights. Calibration techniques can be used to compensate for uneven lighting to generate a uniformly lit image, equivalent to a scanner image.

Older readers, which only had to deal with reading the single OCRB font found in MRZ's, could be ROM based peripheral devices with limited memory and processing power which communicated by RS-232 interface. Today, readers based on a dedicated PC architecture have the advantage of being able to be quickly upgraded with software enhancements, upgraded with faster processing power, vastly superior communication options, and larger memory space which allows processing of multiple high resolution color images.

Due to the fact that the entire field of view can be monitored continuously, the video image feed can itself be used as a sensor to detect when an ID has been inserted in the reader. Upon detection, the image can be located, deskewed and cropped to contain just the ID image. This image can be compared quickly against a knowledge base of known document types. The type of document can be verified by checking for the presence of certain known distinctive features. Given a known document type, then additional images under appropriate lighting conditions can be captured and the layout for that particular document can be obtained from the knowledge base.

Extraction is the process of deriving usable information or images from the document fields. A given area may have very field specific image processing operations applied, such as contrast enhancement, color filtering, dilation/erosion, sharpening, or others. In some cases, this is to enhance images before applying OCR processing to fields such as MRZ's, ID number, Name, Birth Date, or Expiration Date. In many cases, text information is available in the IR image with the colored scenic background dropping out. Using appropriate OCR engines, even passport punched text or non Roman alphabets, can still be read. OCR results can be post-processed, for instance to create a uniform format for dates or perhaps to calculate current age from the birth date. Likewise, barcode images can be decoded from the image itself if the resolution is sufficient.

One of the most useful features of a reader/authenticator system is the ability to recognize arbitrary patterns that may occur in documents. One use of this ability is to test for the presence of certain sometimes subtle features which are

markers for different issues of a given document. Another is for the verification that specialized patterns used for security purposes are present. Multicolored UV patterns are now being commonly used. Sometimes breaks in the patterns may indicate tampering. These types of authentications are critical in today's environment. While a forged document may look almost perfect to the eye, getting all the document elements correctly in all lighting conditions is more difficult to achieve for forgers.

A large library of authentication tests can be developed and used. These tests can be tailored to the types of forgeries or modifications likely to be done a particular type of ID. An authentication test is possible for any of the security features present in a given document. There can be general tests for the presence of certain colors or ink sensitivity in various lighting conditions, for instance to see if there is an IR component to ink or a strong UV component where there should not be. Cross checking between different exemplars of the same information, e.g. the birth date derived from the MRZ and that displayed elsewhere on the passport can be used for authentication. With the ability to decode the barcodes, magnetic stripes, or smart chip info, comparison can easily be made between information derived from these features and those derived by OCR extraction of the text information from the corresponding human readable fields. On passports, the MRZ information is easily compared with information from the upper portion of the document. Certainly, any information garnered from text reading such as name or ID number could be used to query an external data base to verify the validity of the document's other information. Such an inquiry could also return a stored photo or other biometric. More discussion of the various types of security features can be found [5, 9].

The importance of any authentication checks can be weighted. They can be consolidated to arrive at a risk score for a particular ID presenter. The risk can be weighed against the entitlement that presenting the card allows. An automated reader system provides an audit trail of document inspection. Images captured from the process can easily be stored or forwarded for more detailed analysis, not necessarily by the frontline inspector. A networked system could funnel any risk cases for a secondary inspection. In many situations, there are other opportunities to apprehend the presenter of a false ID, e.g. in the case of airport screening, before boarding the plane or even upon disembarking at the destination.

One of the most critical components of an automated authentication system is its knowledge base of document characteristics. It must be secure and encrypted to prevent potential forgers from using this information. By use of an encrypted knowledge base, even the inspectors, who may potentially be subject to compromise, may not be privy to security features that are being used in the authentication process. The knowledge base must be capable of being frequently updated as new documents and variations are issued. Being able to easily train the knowledge base for new documents is an important component. It must be flexible and expandable in its ability to deal with new security features that may be added. It must be adaptable

and programmable in terms of the risk incurred by a given security feature alert (which could be due to dirt or wear). The knowledge base is maintained through cooperative arrangements with government agencies and other access to known good and falsified documents.

Inspectors are presented with a tremendous variety of ID's and have a difficult time authenticating them and keeping up to date with what constitutes a valid document. A new generation of reader/authenticators can help automate the ID inspection process for existing travel documents. These devices are useful in detecting totally forged documents, modifications to otherwise valid documents, and flagging the use of valid ID's by different presenters. Use of these new units minimizes the dependency on the inspector's document expertise, helps them be more efficient, and provides a greater degree of security.

References

- [1] James Hesse, "Counterfeiting and Misuse of the Social Security Card and State and Local Identity Documents," Testimony before U.S. House Judiciary Committee, Subcommittee on Immigration and Claims, July 22, 1999.
- [2] Max Forge, How to Make Driver's Licenses and Other ID on Your Home Computer, Loompanics Unlimited, Port Townsend, Washington, 1999.
- [3] Sheldon Charrett, Secrets of a Back Alley ID Man: Fake ID Construction Techniques of the Underground, Paladin Press, Boulder, Colorado, 2001.
- [4] Gary Thomas, "Bush Administration Asks for Extension of High Tech Passport Deadline," Voice of America News, March 24, 2004.
- [5] American Association of Motor Vehicle Administrators (AAMVA), Personal Identification – AAMVA International Specification – DL/ID Card Design, September 25, 2003.
- [6] International Civil Aviation Organization (ICAO), Document 9303, Machine Readable Travel Documents, Part 1 — Machine Readable Passports (2002), Part 2 — Machine Readable Visas (1994), Part 3 — Size 1 and Size 2 Machine Readable Official Travel Documents (2002), ICAO, Montreal, Quebec, Canada.
- [7] Asa Hutchinson (Under Secretary DHS), Testimony before U.S. Senate Committee on Finance, September 9, 2003.
- [8] Drivers License Guide Company, ID Checking Guide, 2004 Edition, Redwood City, CA, 2004.
- [9] Bruce Monk, "Designing Identity Documents for Automated Screening", 2004 IEEE Conference on Technologies for Homeland Security, Cambridge, MA, April 21-22, 2004.