The author(s) shown below used Federal funds provided by the U.S. Department of Justice and prepared the following final report:

Document Title:	**Technology Transfer of Forensic Document Analysis System: Final Report**
Author:	**Yong Shin**
Document No.:	**231951**
Date Received:	**September 2010**
Award Number:	**2006-DN-BX-K025**

Technology Transfer of Forensic Document Analysis System: Final Report

Submitted by

Cedartech Inc., Williamsville, New York 14221
Tel: 716-983-5834
Yong Shin, Principal Investigator
Email: ycshin@cedartech.com

to

US Department of Justice
Office of Justice Programs
National Institute of Justice
810 Seventh Street N.W.
Washington D.C. 20531

October 18, 2009

Abstract

Several basic studies have been supported in the past by the NIJ on computer algorithms for handwriting analysis so as to support Questioned Document (QD) Examination. Many of the resulting algorithms have also been assembled into a prototype software system. The purpose of this project was to make the software more available and usable to QD examiners for case-work. The principal effort was that of outreach to forensic practitioners in the form of workshops, presentations, and exhibits in trade shows. The project also involved software testing, creating a user manual and making software updates available for download. These efforts have resulted in a more usable system, several licensees for the software, a website from which a significant number of software downloads have been made, a general awareness of the existence of such a tool by the QD community and valuable feedback for further improvement. It is concluded that the tools provided by CEDAR-FOX in its current state can be of use to the forensic community, while at the same time further refinements of many aspects are needed.

Contents

1   Executive Summary

Research has been previously conducted on topics related to handwritten questioned document examination at the University at Buffalo with support provided by the NIJ. Some of the work was on determining the extent of individuality of handwriting, study of the discriminability of the handwriting of twins, algorithms for feature extraction and tools to assist the document examiner. Such research resulted in several algorithms which were then assembled at the university into a prototype software system known as CEDAR-FOX.

The purpose of the present project was to address the last mile of making the software system available to QD examiners. A software prototype developed in a university environment cannot be easily made available to users and hence a concerted effort was made under this project to transfer the technology. The software system, whose intellectual property rights are owned by the university, were licensed by Cedartech to perform the technology transfer function.

Under this project the user manual for the system was improved. A web portal was created and the software made available for download over the Internet. Several new releases of the software were uploaded into the site.

The software system was demonstrated to the Questioned Document (QD) community at national events. These included meetings of ASQDE, SWAFDE, SAFDE, AAFS and IAI. In some of these forums the participants were able to experiment with the software using a hands-on approach to test the capabilities. At others presentations were made explaining the nature of the software.

These efforts resulted in about a half-dozen institutional licenses. Nearly a hundred downloads were made. Requests for download continue to come on a regular basis– about two or three per month from all over the world. Much useful feedback was obtained on improving the system. They included ideas on tools that would be most useful, relative emphasis on project directions and specific suggestions on techniques.

This effort can hopefully serve as a model for other forensic disciplines to bridge the gap between research and practice. Some of the functionalities of CEDAR-FOX are advanced, e.g., it quantifies the strength of evidence on an opinion scale which is still to be accepted by the user community. Recent discussions suggest that all of the impression evidence domains will have to proceed in this direction.

2 Final Report Narrative

2.1 Introduction

2.1.1 Statement of the problem

Basic research on computational methods for handwritten questioned document examination has been conducted at the University at Buffalo with funding provided by the NIJ to the Research Foundation of the State University of New York. Examples of such work are: determining the extent of individuality of handwriting[1], study of the discriminability of the handwriting of twins [2], development of feature extraction algorithms and tools to assist the document examiner. This resulted in several algorithms which were then assembled into prototype software system known as CEDAR-FOX[3].

The only other comparable software available to law enforcement agencies is the FISH system, and its successor called WANDA, which have been developed by German law enforcement agencies. The sustained research effort that led to CEDAR-FOX has functionalities that are more state-of-the-art. The German systems are not available to all organizations. The proposed modest cost of the CEDAR-FOX license was to allow it to be more generally available to police departments and law enforcement agencies.

The goal of this project was to make efforts in transferring a software system for forensic handwriting examination to end-users. Since the QD user community is quite small, such products are not sufficiently commercially viable to make them self-supporting. Although the research developments are still ongoing, what has been developed so far can be of immediate value to the forensic document analysis community.

This outreach effort is distinct, but complementary, to the research conducted on new algorithm development, and supported by the NIJ, at the University at Buffalo. Such a distinct effort was considered appropriate since outreach cannot be effectively performed at the university.

2.2 Methods

The software system CEDAR-FOX has several capabilities commonly encountered in forensic examination. They include: Writer Verification, Writer Identification, Signature Verification, Signature Identification, Document property extraction and Document sub-image retrieval. A screen shot of the top-level functions of the system is included in the Appendix(see Figure 1).

While CEDAR-FOX runs under the Windows operating system it is available to users of other operating systems such as Apple Macintosh due to the ubiquity of Windows emulators such as VMWare.

The software was licensed by Cedartech from the Research Foundation of the State University of New York which owns intellectual property rights to CEDAR-FOX in the form of a recently awarded United States patent[4].

Tasks performed under this project were as follows:

2.2.1 Software Download

The CEDAR-FOX software was made available for easy downloading on the Cedartech website at http://www.cedartech.com. A new web portal for CEDAR-FOX was developed and launched. The download requires a user name and password which were provided liberally to any serious user.

An introductory presentation slides using the PowerPoint template was developed and made available in PDF format from the new CedarTech website. It shows the overview of the software, the computational aspect of the software, and future development plans.

Test versions were made available free of charge. Licensed versions had a fee associated. The current individual software license price is $500.

2.2.2 Software Testing

This involved systematic testing of the software under various scenarios of forensic document analysis, e.g., multiple known writing. The goal was to make sure that the system does not crash under certain unexpected user choices. The sources of software crashes were discovered, e.g., memory leaks, and suggestions were made to reconfigure the previous version of the system.

2.2.3 User Manual

The previously existing preliminary user manual was expanded to anticipate typical system uses and also to have an FAQ section. A screen shot from the user manual is included in the Appendix (Figure 2).

2.2.4 Software Releases

In 2007 CEDAR-FOX version 1.1 was released. It included a new capability of handling compressed images in TIFF image format. With the new capability, the software could read and write compressed TIFF images in addition to PNG images. For the development of the version 1.1, Mr. Chen Huang, a doctoral student at the university, was hired as a summer intern at CedarTech. The release of the version 1.1 was introduced to a number of state and local law enforcement agencies and individuals as well as the private and other agencies and individuals from the United States and other countries.

Version 1.2 was released to reflect some bug fixes such as memory leaks and minor improvements as well as the newly completed interface for compressed TIFF images.

Version 1.3 has steadily been used since 2008 while bug fixes continue to be made. Potential licensed users found some bugs in signature verification modules and suggested ideas for better user interface. The suggestions have been collected and forwarded to the university research group for future releases.

2.3   Results

2.3.1   Software Licenses

During 2008 there were four paid licensees: The Netherlands Forensic Institute, Georgia Bureau of Investigation, Canadian Public Safety and Canada Border Service Agency. During 2009 a site license was acquired by West Virginia university for inclusion in their forensic curriculum.

Nearly a hundred downloads of the software have been made across the globe. Many request come each month to download the software. Some of these are for instructional purpose.

2.3.2   User Feedback

Valuable feedback has been received to improve software performance. They include not only identifying software bugs but also establishing the relative importance of tasks, e.g., signature verification versus comparison of extended writing, ability to hone in on features that are distinctive, etc. For instance the importance of characters that could not be easily recognized are important in that their unique nature makes them useful for writer identification (see Figure 4) . Such feedback has been passed along to the university team to further improve the software.

2.4   Conclusions

2.4.1   Implications for Policy and Practice

This project had the primary goal of technology transfer. Technology transfer of forensic research is a two-way street: while practitioners need to be informed of the results of research, researchers needs to be informed of the type of problems to be investigated. While CEDAR-FOX has many useful tools for the practitioner, the complexity of the software makes it necessary for user training for greater acceptance.

The software provides results on an opinion scale which has to be accepted by the community particularly in light of courtroom testimony moving in the direction of accepting and quantifying uncertainty. CEDAR-FOX provides an example of how such quantification can be made and which could be adopted by other forensic disciplines.

2.4.2   Implications for Further Work

As in the case of other areas of impression and pattern evidence there is much more to be accomplished in automated tools for handwriting examination. There is also much more fundamental research that needs to be accomplished, such as document examiner suggested features, improved methods for signature verification and more robust methods for quantifying the strength of evidence.

3   Dissemination

One of the main objectives of this NIJ-funded effort was the dissemination of information to the forensic community. Several different efforts were made as described

below.

3.1   Mailing List

The availability of the software was made known to the QD community through mailings and publicity at forensics meetings. A list of potential users was compiled using the information obtained from multiple sources. Included in the list were users who requested trial version, and attendee list from the 2007 Applied Technologies Conference. Also, participants who attended earlier AFDE Symposium and ASQDE Conference were included in the list. The list was used in announcing the new release of CEDAR-FOX. The list includes a total of 260 entries, including 189 state and local law enforcement agencies/individuals and 71 private and other organizations/individuals.

3.2   Trade Shows

3.2.1   IAI 2007

A booth was set up at the IAI Conference in San Diego, CA during July 2007.

Brochures and CD-ROMs containing the trial version of CEDAR-FOX were created. Also a banner was newly designed. CedarTech participated in this conference by sending two exhibitors (Mr. Chen Huang and Mr. Harish Srinivasan). They were able to present CEDAR-FOX to the visitors at the booth during the two-day exhibition. Prior to the trip, all the presentation materials have been prepared or secured projector screen with tripod, projector, notebook computers, etc. They arrived the day before the exhibit open, and moved out the day after.

About 30 people stopped by the booth. Two university people expressed interest in acquiring the software system for teaching purposes. Several other new ideas were generated for CEDAR-FOX by visiting various booths. Follow-up email messages were sent out.

3.2.2   ASQDE 2007

At the Annual Meeting of ASQDE in Boulder, CO during August 2007, CedarTech had a booth with two exhibitors (Dr. Yong Shin and Mr. Chen Huang). The booth included presentation materials, brochures, CD-ROMs, computers, banners and other equipment.

At a conference plenary meeting on the first day during opening remarks, Dr. Shin introduced CEDAR-FOX using a slide presentation.

3.3   Magazine Advertisement

In order to widely inform the law enforcement and forensic communities the availability of CEDAR-FOX software, advertisements were placed in Evidence Technology Magazine and Forensic Magazine in the September/October 2008 issue and December 2008/January 2009 issues. The ads introduced CEDAR-FOX with an itemized description including key features such as the capabilities of handwriting identification, writer

and signature verification, handwriting segmentation, and searching scanned images without OCR.

3.4 Workshop

A one-day invited workshop on using CEDAR-FOX was held at the annual meeting of SWAFDE (Southwest Association of Forensic Document Examiners) on April 11, 2008. This was conduced at the expense of SWAFDE. There were about 40 attendees at this all-day event which provided many useful ideas for further improvements of the software

3.5 Conference Presentations

3.5.1 AAFS 2008

A presentation on CEDAR-FOX was made at the NIJ Grantees meeting held in conjunction with the AAFS meeting in Washington DC on February 19, 2008.

3.5.2 AAFS 2009

CEDAR-FOX was presented in QD section of the annual meeting of AAFS in Denver, CO in February 2009. About 30 CDs were distributed to attendees.

3.5.3 SAFDE 2009

CEDAR-FOX was presented to the Southern Association of Forensic Document Examiners (SAFDE) at their annual meeting in Atlanta, GA in April 2009. About 20 CDs were distributed to the attendees.

4 References

The following publications are referred to in the narrative.

References

[1] Srihari, S.N., Cha, S., Arora, H., Lee, S.: Individuality of Handwriting. Journal of Forensic Sciences **47** (2002) 856–872

[2] Srihari, S.N., Huang, C., Srinivasan, H.: On the the discriminability of the handwriting of twins. Journal of Forensic Sciences **53** (2008) 430–446

[3] Srihari, S.N., Srinivasan, H., Desai, K.: Questioned Document Examination using CEDAR-FOX. Journal of Forensic Document Examination **18** (2007) 1–20

[4] Srihari, S.N., Shin, Y.C., et.al.: Method and Apparatus for Analyzing and/or Comparing Handwritten and/or Biometric Samples (2009) United States Patent No. 7,580,551.

5 Appendix: CEDAR-FOX screen shots

In the following a few screen shots of CEDAR-FOX are provided so as to indicate the extensive nature of tools available and of the user manual, which was of particular interest in this project.
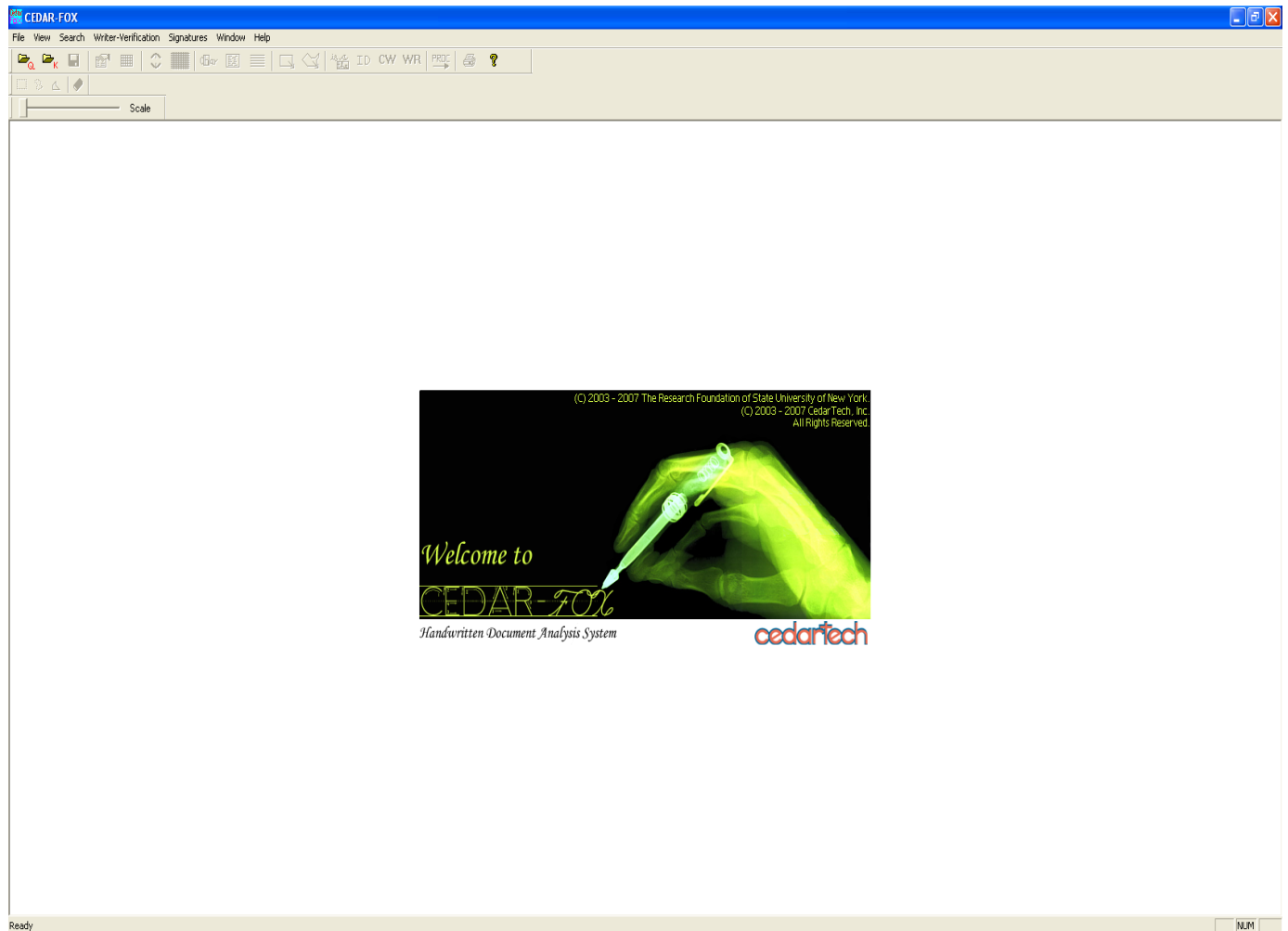
List of Figures

Figure 1: Main menu of CEDAR-FOX system. The pull-down menus on top are: File, View, Search, Writer-Verification, Signatures, Window and Help. Clicking on the Help button provides a pull-down menu that contains an entry for the User Manual
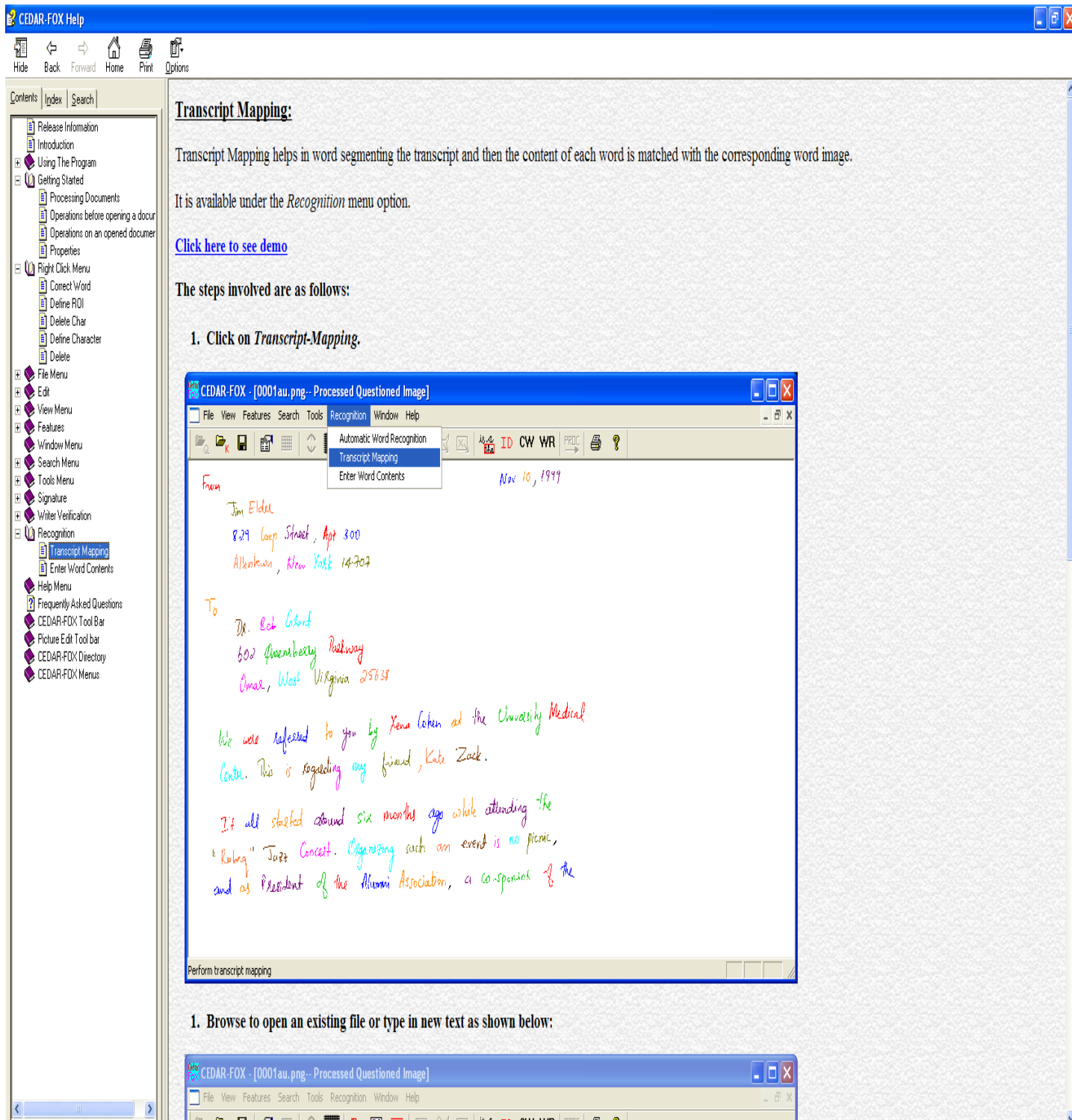
Figure 2: A sample User Manual Page. The panel on the left lists Help topics. Text and figures on the right explain the topic, which in this case is about the Transcript Mapping function which allows the user to entry the word truth for each word in the input document.
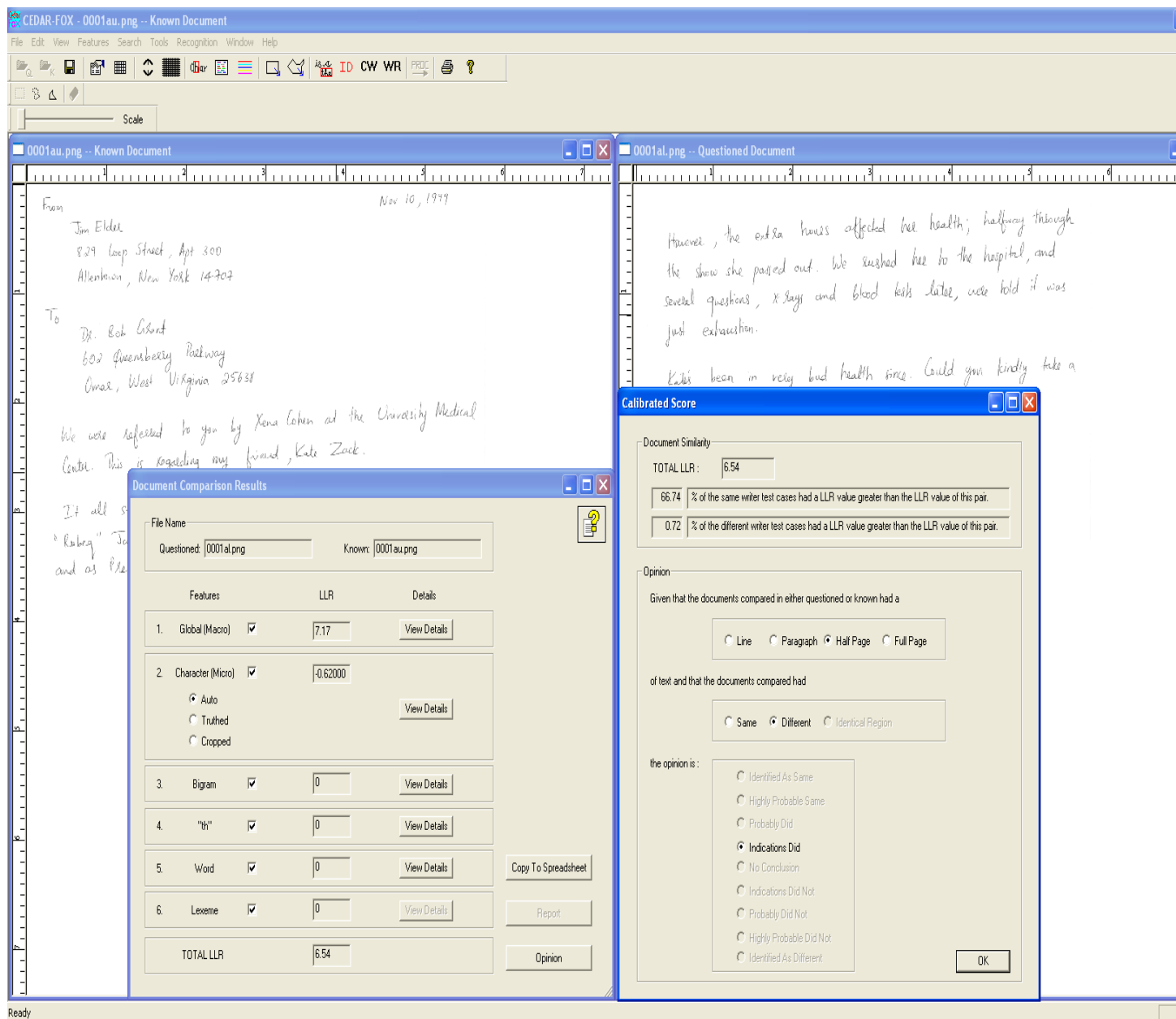
Figure 3: Opinion on Comparison of Two Writing Samples. The results of the comparison on two handwriting specimens are displayed in two pop-up windows at the bottom of the screen. The one on the left gives the log-likelihood ratio of five feature types and of the total. The one on the right converts this score into a nine-point opinion scale. In this case the score was 6.54 which maps into a weak indication that it was written by the same individual.
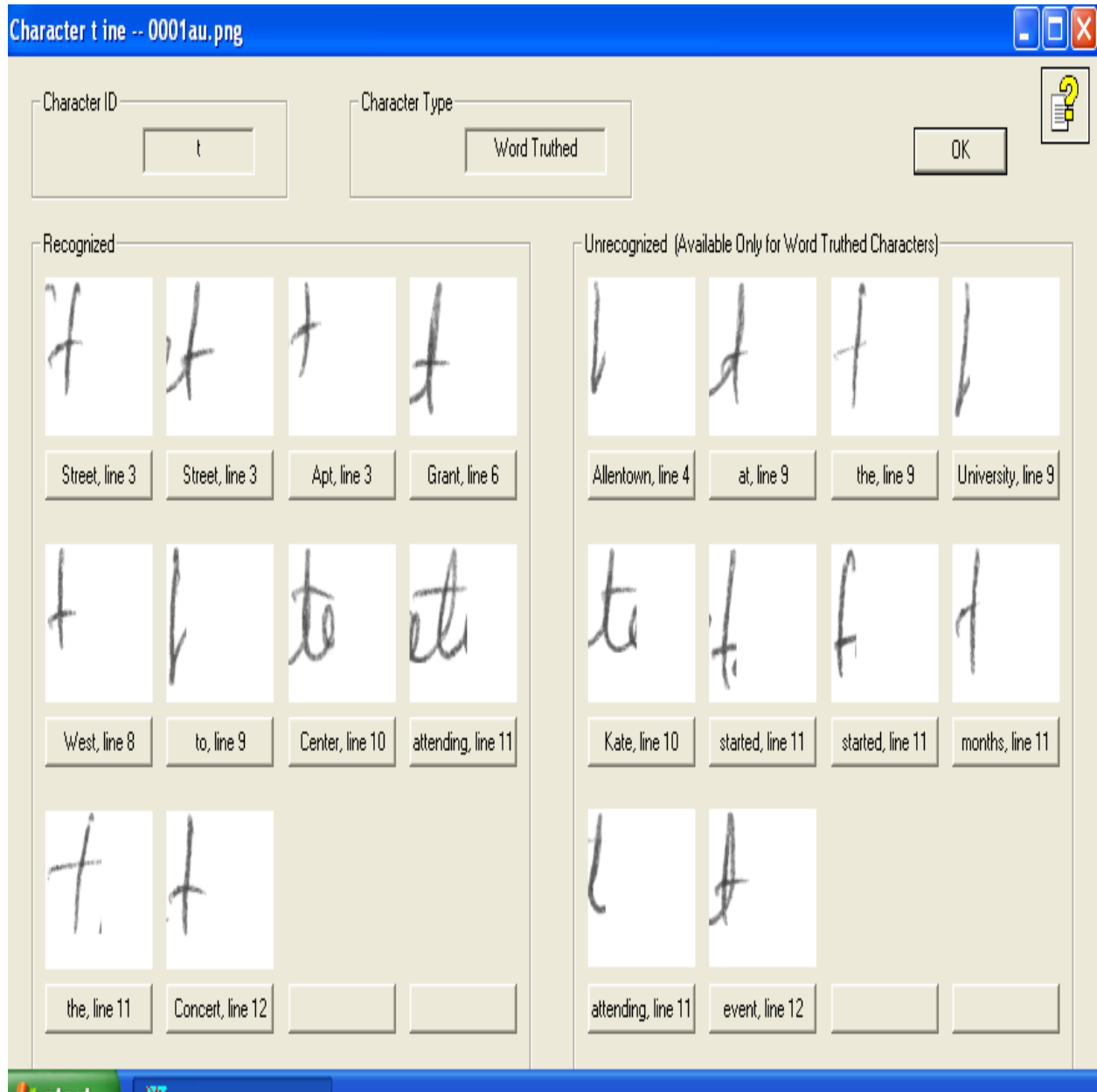
Figure 4: A particular letter 't' pulled out of a document. Such instances are useful to a document examiner in manually comparing documents.

Figure 5: Word Search. The user can type in a search string for which the system will retrieve postential matches in the handwritten document.
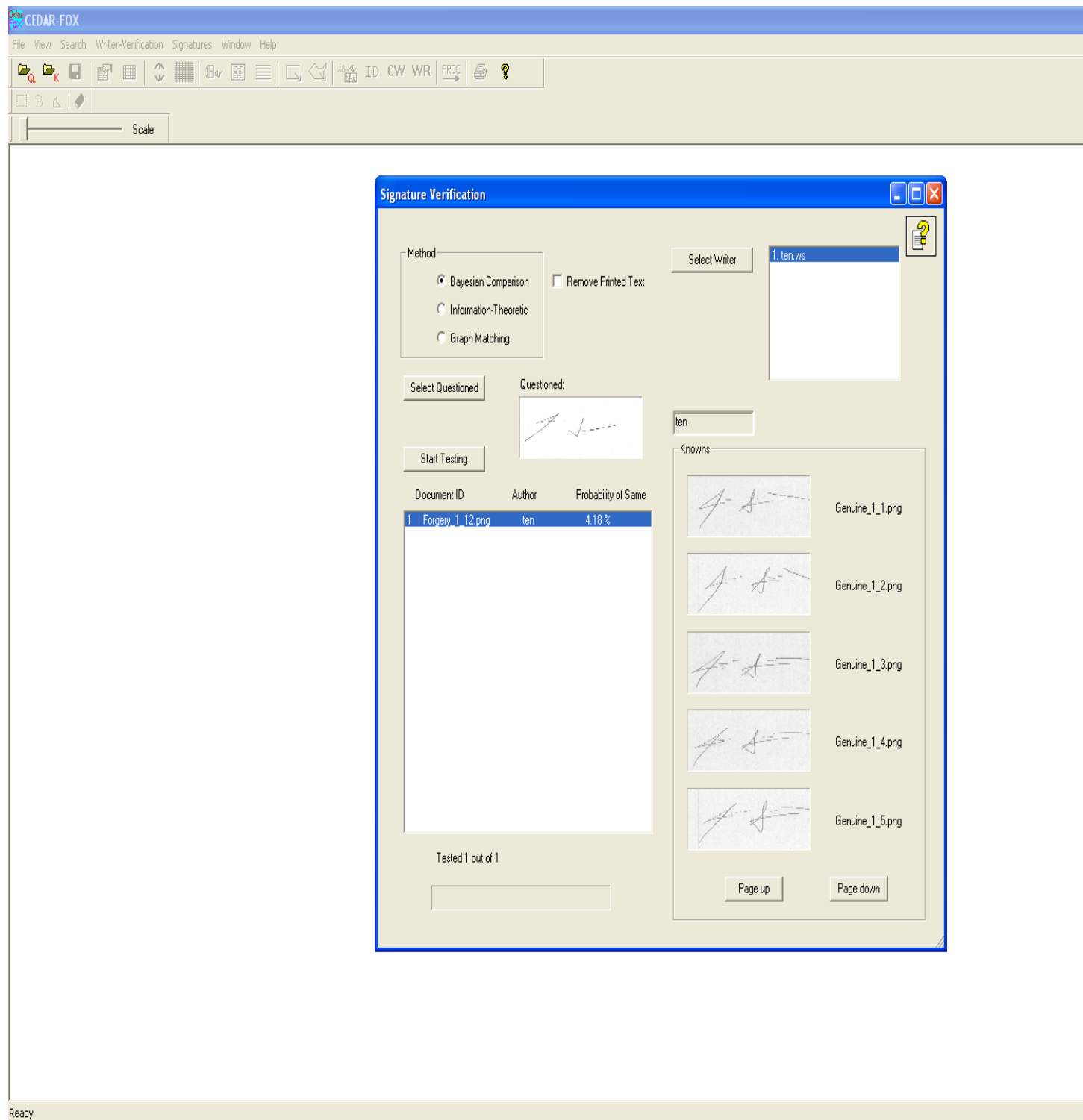
Figure 6: Result of Signature Verification. A questioned signature in the middle of the screen is compared to known signatures in the right to provide a confidence of the signature belonging to the known set as 4.18% or being a forgery.