NIJ Fellowship Applications

Amy Crawford, Nate Garton, and Kiegan Rice

April 11, 2018

Application Process

Description of Solicitation

Title: Graduate Research Fellowship in Science, Technology, Engineering, and Mathematics

- ▶ Work must have demonstrable implications for addressing the implications of preventing or controlling crime, and/or the fair and impartial administration of criminal justice in the U.S.
- ▶ Areas of interest:
    - ▶ Reducing crime (particularly violent)
    - ▶ Protecting police officers and other peeps
    - ▶ Issues concerning the opioid abuse epidemic
    - ▶ Victimization (human trafficking)
    - ▶ "Supporting prosecutors in their efforts to meet their mission"
    - ▶ Illegal immigration issues
- ▶ Those considering forensic evidence research should look at:
    - ▶ OSAC Research Needs
    - ▶ NIJ Technology Working Group list of research areas
    - ▶ NIJ Core Science and Technology Research Objectives

Why we were chosen to apply (student status)

- Up to 3 years of funding for a dissertation
- Eligibility: Enrolled in a doctoral STEM program, proposal of dissertation that is relevant
- Why us?
  - Literally, WHY US?
- Early in the process of dissertation research
  - Doesn't help to apply if you're almost done
  - All had a vague idea of a dissertation that seemed to fit the solicitation

How Amy's research fits in

Title: A Novel Application of Machine Learning Methods: Writership and Complexity in Forensic Handwriting - Handwriting feature extraction and selection - Inter- vs intra-writer variability analysis - Complexity analysis (unsupervised learning) - Similarity score and construction of reference distributions (supervised learning) - Provide an online tool for interested parties

How Nate's research fits in

Title: Spatio-temporal point processes for crime (STOPPR)

- Crime modeling and prediction
- Bayesian spatio-temporal point process models
- Provide a framework and hopefully a tool for others (criminologists, law enforcement) to make predictions or test hypotheses

How Kiegan's research fits in

Title: Strengthening foundational validity of 3D imaging in bullet examinations: persistence and variability of scans

- ▶ Secondary Analysis of Striation Persistence Data
- ▶ High-Resolution Microscopy Variability Study
- ▶ Comparison of several currently proposed methods for analysis
- ▶ Adding more information to the world of 3D bullet imaging
- ▶ Testing out sensitivity of methods on new/different data!

The Process

- Many documents that needed to be prepared:
  - Budget detail / narrative (Marc and Stacy prepared these)
  - Conflict of interest form (template)
  - Project Abstract (400 words)
  - Statement of Support from Committee Chair (thanks everyone!)
  - Undergraduate Transcripts (WHY...?)
  - Graduate Transcripts
  - Enrollment Verification
  - **Research Narrative AND APPENDICES**
    - Bibliography/References (supposed to be fairly comprehensive)
    - Curriculum Vitae/Resumes (of student and advisors)
    - Personal Statement (2 pages, including career goals)
    - List of dissertation committee (template)
    - Proposed timeline/milestones (we honestly have no idea)
    - Privacy Certificate (weird form)
    - Letters of Cooperation from outside collaborators (Thanks Gary and Vic!)

Research Narrative

- ▶ Structure:
  - ▶ 12 pages max (minus title page, contents)
  - ▶ Title page (surprisingly complicated)
  - ▶ Table of contents (easy enough)
  - ▶ Statement of Problem and Research Questions
  - ▶ Project Design and Implementation
  - ▶ Capabilities and Competencies
- ▶ Things to remember:
  - ▶ Research need in area of study
  - ▶ Current gaps in data, research, and knowledge
  - ▶ Discuss previous research relevant to the problem
  - ▶ Data acquisition methods (in detail)
  - ▶ Demonstrate validity and relevance of data to be collected
  - ▶ Justify methods of data analysis
  - ▶ Address feasibility and speculate potential challenges, plans to mitigate them
  - ▶ Plans to make results available to interested parties
  - ▶ Capabilities of the student and the advisor
  - ▶ Academic environment and supporting resources
  - ▶ Project management plan

Pros of the Process

Pros

- ▶ We had a lot of help!! (Thanks Stacy, Sarah, Marc, and Harlie!)
- ▶ Forced us to form a research plan
    - ▶ What the research questions are
    - ▶ How we are going to address the questions we have
- ▶ Gave us each a semi?-comprehensive lit review (base for going forward)
- ▶ Now we all have these materials ready to work off of moving forward
- ▶ Know what the process looks like
- ▶ Would be really good for CSAFE as an organization
    - ▶ Expanding on current research
    - ▶ Adding a cool new type of research to the pot

Cons of the Process

Cons

- HUGE amount of time and energy - developing research narrative
- HUGE amount of time and energy - all the appendices/documents
    - Big organizational challenge
    - Large group of people involved - gets messy!
- Short notice
- Lack of familiarity with the process

Outline of Research Narratives

Kiegan: Bullet Data

How did I decide on my research questions?

- ▶ Have been working with bullet data
- ▶ Automated methods for groove identification in 3D bullet land scans
- ▶ Learning more about the current state of research at conferences, etc.
- ▶ Some interest in 'relevant populations', and doing comparisons with representative data to back it up.

Background & Literature

- Comparison of bullet striations
- Issues with lack of foundational validity
- NRC, PCAST reports
- Some initial models (Chu et. al. at NIST, and CSAFE)
  - Cross-Correlation Functions, QCMS, Random Forest
- Initial persistence studies (Bachrach)
  - Data unavailable

Research Needs

NIJ Technology Working Group

- Fundamental understanding of how environmental factors can affect evidence
- Time, scanning process
- Scientific foundations for the evaluation of evidence in support of qualified and definitive conclusions
- Support for standards development and validation of methods

OSAC - Whether QCMS withstands the transfer from 2D to 3D

Research Questions

1. How comprehensive and conclusive are currently available data on persistence of striae, and what additional data need to be collected to fill informational gaps?
2. What amount and sources of variability are introduced by the 3D scanning process; in particular, how are 3D scans of bullet lands affected by differences in microscope and operator for different brands and calibers of gun?
3. What is the impact of variability in the 3D scanning process and differing brand-caliber combinations on accuracy and precision of proposed methods for automated comparison of bullets?

Proposed Studies (Data Collection)

1. Groove Identification (they are getting this paper for "free")
2. Secondary Analysis of Striation Persistence Data
   - Identify gaps in data that need to be filled
   - Differences in persistence across different types of gun?
3. High-Resolution Microscopy Variability Study
   - Gauge Reproducibility and Reliability (Gauge R&R)
   - Repetition of scans for operator, machine, day

Proposed Studies (Data Analysis)

4. Sensitivity of Automated Methods
   - Taking collected data "grid"
   - Running through several proposed algorithms
     - Eric's Random Forest
     - Chu (NIST) Cross-correlation function
     - Chu (NIST) Quantitative Consecutively Matching Striae
   - Testing whether accuracy changes based on differences in bullet

Dissemination of Research

- Journal of Forensic Sciences, Annals of Applied Statistics
- AFTE, AAFS Meetings
- All collected data made publicly available through NIST
- Proposed timeline is semester-by-semester

Amy: Handwriting

Contents

Current Gaps/Research Needs

Research needs identified by the NIJ TWG

Technology Working Group Operational Requirements Updated February 2018	Scientific Research	Technology Development	Policy or Protocol Development	Assessment & Evaluation	Dissemination &/or Training	Other	Forensic Discipline
. . .							
Scientific foundations for the evaluation of forensic evidence in support of expert conclusions	X	X				X	Impression & Pattern/ Trace Evidence
Determination of accuracy and reliability of forensic methods and conclusions, including potential sources of error	X	X					Impression & Pattern/ Trace Evidence
Novel statistical approaches for the interpretation of forensic evidence	X					X	Impression & Pattern/ Trace Evidence

Current Gaps/Research Needs

Research needs identified by the NIJ TWG

Technology Working Group Operational Requirements *Updated February 2018*	Scientific Research	Technology Development	Policy or Protocol Development	Assessment & Evaluation	Dissemination &/or Training	Other	Forensic Discipline
⋮							
Reference collection databases of handwriting samples, copybook curricula, and typewriter and computer font styles						X	Forensic Document Examination
Quantitative assessment of intra- and inter-person handwriting and handprinting variation	X	X		X		X	Forensic Document Examination
Assessment of the comparability of different forms of writing from individuals (e.g., initials, signatures, handwriting, hand printing, foreign writing)	X						Forensic Document Examination

Current Gaps/Research Needs

Research needs identified by the OSAC:

Current Gaps/Research Needs

Research needs identified by the OSAC:

Forensic Document Examination:

Comparability and Complexity in Handwriting

Hand Printing Complexity and Comparability

National Database of Handwriting

Validation of Conclusion Scale

Current Gaps/Research Needs

Research needs identified by the OSAC

OSAC RESEARCH NEEDS ASSESSMENT FORM

Title of research need: Comparability and Complexity in Handwriting

[. . . background information, references, etc. . . .]

4. Status assessment (I, II, III, or IV):   I

	Major gap in current knowledge	Minor gap in current knowledge
No or limited current research is being conducted	I	III
Existing current research is being conducted	II	IV

This research need has been identified by one or more subcommittees of OSAC and is being provided as an informational resource to the community.

Objectives

1. Following construction of a **handwriting dataset** that will be publicly available and will support both research and case work, **extract and identify a set of features** that have high discriminating power.

Objectives

2. Conduct a statistical analysis of **complexity and comparability** of written samples.

Objectives

3. Develop a statistical modeling approach to **combine features into a single similarity score** that can be used to compare two handwriting samples.

▶ Once we have chosen a method with potential, we will validate the algorithm to the extent possible with the writing samples available.

Objectives

4. **Assemble distributions of similarity scores** among writing samples known to have been produced by the same individual and writing samples known to have been produced by different individuals.

Objectives

As part of the development we will...

- ▶ Assess importance of features (individualizing characteristics), writing complexity, and the relationship between the two.
- ▶ Characterize statistical inter- and intra-writer variability at the level of individual features and also at the level of similarity scores.
- ▶ Quantify error rates.

On to Project Design

Contents

Feature Extraction and Selection

- Based in the *graphical structure* of the writing
- Other features will be introduced
  - Gradient, Structural, and Concavity (GSC) feature vectors
  - Character recognition
  - Vertical and horizontal projections

- Features for complexity analysis
- **In the end:** combine to a single feature vector for two documents
- Feature selection: parsimony - can we be accurate *enough* while remaining simple *enough* to be approachable and interpretable?

Bayesian Analysis of Writership (not focused on in the narrative)

- ► What I'm up to now
- ► Addresses questions of writership in a closed set and, hopefully at some point, an open set

Complexity Analysis

- (small)
- Examiners have large amounts of variability when assessing the complexity of a peice of writing/signature (Hal)
- **Features**: characterize the complexity of a writing sample
- Unsupervised learning methods

Similarity Scores

- **Features**: characterize the similarities and differences between **two** writing samples
- Supervised learning methods (expand?)

Create Reference Distributions

- Produce pairwise scores from mated and non-mated documents
- Investigate distributions of the scores from mated and non-mated documents

Deliverables

- Papers
- Conferences presentations (AAFS, The American Society of Questioned Document Examiners - ASQDE, others)
- Data
- (Ideal) Software tool to process documents, compute score, and location wrt the mated and non-mated distributions.