ISTeC Retreat on the Scientific and Engineering Foundations of Information Science and Technology

Saturday, May 8, 8:15am to 3:00pm, Lory Student Center

Report on Imaging, Tracking and Computing

Participants

Chair: Michael Kirby (Math, CS)
Participants: Beveridge (CS), Dangelmayr (Math), Draper (CS), Iyer (Stat), Kirby (Math), Kley (Math), DeMiranda (Education), Scharf (ECE), Troup (Psychology), Wang (Stat)

Session Overview

The participants in the breakout session formally entitled Imaging, Tracking and Computing have departmental affiliations in Computer Science, Education, Electrical and Computer Engineering, Mathematics, Psychology and Statistics. We were charged to identify areas of common interest and to both describe tools we have developed and problems that require tools that need to be developed all of which roughly fit under the session theme. Given the diverse yet coherent make-up of the session participants we pursued focused discussions on areas of related interest that included the geometry of information in large data sets, cognitive modeling and biologically motivated algorithms for computation. Specific applications of interest included computer vision and biometrics as well as dynamical systems and neuronal modeling. In addition we discussed several issues related to education and the relationship of the teaching and learning with large-scale interdisciplinary university initiatives.

Open Problems of Interest

A unifying theme of the open problems discussed in this session revolved around modeling information. For example, how does one proceed from data to the understanding of phenomena? How can information be processed to isolate the essential and illuminating features of interest? Hard problems of national interest such as biometrics, e.g., face recognition, and counter terrorism may be viewed, in general terms at least, as problems in the processing of large amounts of coherent information.
In an abstract sense, a central aspect of knowledge discovery in general is the formulation of explanations based the processing and ultimately the understanding of large amounts of data. At the center of this exercise is model development.

General aspects of the modeling process were discussed including

- Model research followed by development of high-speed implementation.
- Given an array of models how is intelligent model selection performed?
- How can the geometry of observed data be exploited in the modeling process?
- Performance bounds.

Potential Scientific Themes for an Institute for Information Sciences

We now summarize the research themes discussed by our session emphasizing the most developed area. Note that an additional theme, i.e., *Dynamical systems and Synchrony of Cell Systems*, is presented as an action task for fostering new collaborations. See the section on Collaborations for more details.

We organize our summary around the areas of significant expertise and national interest as reflected by the session members. Broadly speaking they consist of

- **Biometrics**

and

- **Geometry of Information.**

We argue that these may be viewed as essentially the same topic but perceived differently by the major disciplines involved. We begin by describing one of the central problems in Biometrics, i.e., face recognition, and argue that this may be viewed as a problem in geometry and computation. Then we outline mathematical ideas that makeup the backbone of an emerging field in information sciences, i.e., the geometry of information. These two fields taken together represent a unique array of expertise and problem solving potential.

Biometrics

The field of Biometrics involves quantifying an individual's physical attributes to ascertain their identity. Given we seek to identify an individual, mathematically (and statistically) we must solve the classification problem. Given the expertise at Colorado State University (Beveridge, Draper and Kirby) we propose face recognition as a

paradigm problem for classification. Additional areas of interest include iris and gait recognition.

Classification of faces may proceed once the images are acquired in the form of matrices. For example, a black and white image of a face typically consists of a 500 x 500 array of one byte integers. The goal of face recognition research is to develop algorithms that will identify images of individuals taken under novel circumstances.

Professors Draper and Beveridge in the Computer Science Department and Professor Goevens in Statistics are world leaders in the empirical evaluation of computer face recognition algorithms. Supported by the HumanID project they have developed a state-of-the-art platform for evaluating new algorithms. Additionally, some of the leading research in identifying useful features for recognition known as subject covariates has been carried out at Colorado State by Beveridge, Draper and Goevens.

Professor Kirby has proposed new algorithms to achieve the low-dimensional representations of images of faces and has been a proponent of developing new geometrically based algorithms for classification. In particular, a group in mathematics (Kirby, Kley and Peterson) working in conjunction with Computer Science (Beveridge and Draper) is developing new algorithms for based on algebro-geometric techniques. Professor's Kley and Peterson bring a substantial expertise in algebraic geometry, the tools from which are just beginning to be exploited for algebro-geometric data analysis.

Although his work has been in the context of signal processing, Professor Scharf of ECE and Statistics has developed well-known algorithms for data reduction.

Professor Dangelmayr's primary area of expertise is dynamical systems. He has also explored (with Professor Kirby) the low-dimensional representation of sequences of images, in particular applied to the problem of machine lip reading.

Additionally, the session group possesses considerable in the development of families of models, data detection and estimation, performance bounding, general techniques for dimensionality reduction, geometrical structure of models. expertise [in addition to the above we note Maciejeswki (ECE), Wang (Stat), Iyer (Stat)].

We note that the algorithms developed in the context of the face recognition problem have significant broader impact and can be applied to (for example):

- Classification for failure prediction
- Automatic Target Recognition (ATR)
- Medical diagnosis
- Hurricane tracking
- Target tracking
- Brain interface
    - –Impacts modular Hardware /Software design
    - –Connects to psychology & MCIN

Geometry of Information

To illustrate that the mathematicians at Colorado State University are natural (if not obvious) collaborators on the biometrics problem, we revisit the face recognition problem from a geometric viewpoint.

At the heart of this discussion is the geometry of the raw data. In general, algorithms are not designed to explicitly exploit the fact that, e.g., data may form a manifold rather than a subspace. The example was made that faces do not reside in a subspace and therefore subspace methods may be less than optimal. Indeed, preliminary evidence strongly suggests that nonlinear theory based on geometrical considerations have enormous potential for the development of new classification algorithms.

In view of this, techniques from geometry, differential and algebraic geometry as well as topology have enormous potential to complement statistical approaches for data processing, including biometric data. The emerging collaborations between Computer Science and Mathematics at CSU are motivated by these ideas.

Summary of topics of mutual interest:

•Why should signals and system and data have a low-dimensional underlying geometry structure?
•Algebro-geometric methods [Kley, Peterson, Miranda, Kirby]
•Computation geometry [Peterson, Kirby]
•Large data sets [All]
•Performance bounds and calculations [Scharf, Iyer, Wang]
•Dimension and dimensionality reduction [Wang, Iyer, Scharf, Kirby, Dangelmayr, Miranda]
•Linear vs nonlinear dimensionality reduction [Kirby, scharf, Azimi]
•Class on manifolds (face recognition, etc) [Kirby, Draper, Beveridge]
•Info theory of geometry structures; Kulback-Leibler [Iyer, Wang, Scharf, Chong]
•Geometrical optimization and Fast algorithms [Chong, Scharf, Peterson, Kley, Kirby]

Collaborations

The basic concept of developing and supporting collaborations across disciplines at a single university is generally supported by the members of this session. Indeed, we will provide brief examples (this is an incomplete list) of cutting edge research currently being carried out by members of this session across disciplines.

Project Title: New Tools for Algebro-Geometric Data Analysis, Michael Kirby (Mathematics), Holger Kley (Mathematics), Chris Peterson (Mathematics), Charles Anderson (Computer Science), Ross Beveridge (Computer Science), Bruce Draper (Computer Science). This is a pending project that aims to develop new algorithms based on mathematical theory from algebraic geometry. The intent of the NSF is to support new and especially novel collaborations between Computer Scientists and Mathematicians.

Project Title: Geometric Pattern Analysis and Mental Task Design for a Brain-Computer Interface, Charles Anderson (Computer Science), Michael Kirby, (Mathematics). The primary objective of this project is the classification of multivariate time series associated with EEG patterns via geometric modeling ideas.

Project Title: Understanding Non-Contact Biometrics, J. Ross Beveridge (Computer Science), Bruce Draper (Computer Science), Kevin Bowyer (Notre Dame), Geof Givens (Statistics), Michael Kirby, Holger Kley, (Mathematics), Chris Peterson (Mathematics), Alice O'Toole (University of Dallas) This is a large project intended to develop and evaluate new mathematically based face recognition algorithms.
(This is a pending project).

Project Title: Robust Reinforcement Learning for Control of Heating and Cooling of Buildings, Charles Anderson (Computer Science) and Peter Young (ECE).
Project Title: Interpreting Three Dimensional Images of Growing Plants, Ross Beveridge (Computer Science), Bruce Draper (Computer Science), June Medford (Biology)

Project Title: Analysis of transport, mixing and coherent structures in hurricane formation and intensity, Mike Montgomery (Atmospheric Sciences), Gerhard Dangelmayr (Mathematics), Michael Kirby (Mathematics)

Based on the revelations of areas of common interest in our session we have also proposed a new area for collaboration:

Dynamical systems and Synchrony of Cell Systems

- Connections to dense networks?

- Parallel and distributed information processing [Drape (Math), Troup (Psych), Dangelmayr]
- Modeling neuro-systems, from cells to nets [Troup (Psych), Dudek, Dangelmayer (Math), Anderson, Volbrecht, Neger]
- Control and learning (Dangelmayr)
- Behavior (Dangelmayr, Troup)
- Bifurcation analysis (Danglemayr, Allgower, Tavener)
- How does this basic science impact the engineering of vision systems?
- Learn brain function from the study of dynamical systems that model the dense cellular assemblies. This is a sub-topic in biometric. [Troup (Psych), Dudek, Dangelmayer (Math), Anderson, Volbrecht, Neger]

Action Item: Working group to develop dynamical systems and synchrony of cell systems.

Education

The vision for interdisciplinary collaboration organized around clusters of faculty with complementary expertise and research interests must be expanded to include the education and training of graduate student researchers. Consistent with the educational mission of Colorado State University, graduate student research assistants are an integral part of the research engine at Colorado State University, as they are at all research universities. In the case of interdisciplinary research teams participating students often receive unique benefits stemming from the novel and basic synergistic nature of the activity. Any significant plan for new research initiatives must build into it this assumption and address the infrastructure support for graduate students.

Thus, developing mechanisms for recruiting and supporting highly motivated graduate students should be a high priority and might include:

- Endowed graduate research assistantships
- Graduate student recruitment support via Institute
- Tuition waivers for institute affiliated graduate research assistantships
- Graduate student research awards

It was observed that the number of graduate research assistantships in many of the represented disciplines is significantly lower than desirable but with the positive aspect that this is and area that may afford tremendous growth potential. For example, in mathematics we support approximately 5 graduate research assistantships but with sufficient support we could realistically target mentoring 20 graduate research assistantships.

An institutional commitment to excellence in graduate education not only will greatly assist in the success of potential multi-disciplinary collaborations, it provides a strong evidence to the citizens of Colorado that Colorado State University is performing a critical and essential service to the State.

It was observed that companies such as Cisco that have traditionally supported institutions via technology recognize the importance of investing directly in the workforce. Supporting education is viewed as a long-term investment while lifetime of technology investments can be quite short, e.g., computer systems.

Funding Sources

All the mainstream federal funding agencies support work the session themes: NSF, DARPA, ARO, ONR, and AFOSR. We note that recent increase in funding for Homeland Security and the new collaborations between the Intelligence Community (IC) and the NSF for funding of approaches to combat terrorism.

However, traditional funding sources substantially favor limited collaborations of two or three faculty and possibly one or two disciplines. Attempting to seek funding for a broad program such as that described here may require the development of a funding consortium consisting of several funding agencies.

It is noteworthy that several smaller grants (200-800K) and even some larger grants (2-3 million) involving participants in this session are currently either funded or are pending.

A good overview of particular funding agencies may be found by considering many of the current and pending grants of the session participants.