Research I Foundation

Annual Report

2010-2011

Contents

Research Activities of PhD Students

Aditya Nigam

Annual Report 2010-11

Comprehensive Examination

I have cleared my comprehensive examination on the topic **"Automatic Segmentation of Human Brain Tractography Data"**. Any brain fiber is represented as a set of points in 3D space (typically 20 to 30 points per fiber). The Tractography technique produces thousands of fiber trajectories per subject (nearly 250K). This Tractography data can be seen as a 3D point cloud but that is not very useful. The useful information can be extracted only when they are organized into anatomically meaningful structure.

The problem statement was *"Given the tractography data of a human brain, segment it automatically into tracts having similar fibers which are anatomically meaningful"*.

State of The Art seminar

In the next semester I was preparing for my SOTA (State of The Art) seminar. I have given my SOTA seminar on **"Use of Physiological Characteristics for Personal Authentication"** at Mon, 25 Apr 2011. In my State-of-The-Art seminar I have presented several personal authentication systems (PAS), primarily extracting textures for personal authentication. Some of them are as follows:

1. ***Palmprint:*** Techniques required for Palmprint extraction and enhancement were discussed. Several local and global feature extraction technique such as gabor filtering, phase difference (DFT), discrete cosine transform(DCT) along with some fusion techniques were presented. Matching is mostly done using hamming distance.

2. ***Iris:*** Techniques required for Iris extraction and enhancement were discussed. Several feature extraction technique such as gabor filtering, phase only correlation (POC) and band limited phase only correlation (BLPOC) along with fusing gabor and BLPOC techniques were presented. Here, also matching is mostly done using hamming distance.

3. ***Knuckleprint:*** Techniques required for knuckleprint extraction and enhancement were discussed. Several local and global feature extraction technique such as gabor filtering, phase only correlation (POC) and band limited phase only correlation (BLPOC) along with fusing gabor and BLPOC *etc* are discussed. Matching is mostly done using hamming distance.

Thesis

We have started to work in face recognition in varying environment. Feature extraction based on hausdroff distance [1] is applied giving good results. Later, new edge based weighing technique [2] is applied to get better results.

Recently we have started to work on texture based feature extraction as in most of the traits texture patterns are observed and are very discriminative. In the mean time I am reading about techniques such

as DFT, DCT, Gabor, Corners, SIFT and SURF features and their applications in traits like palmprint, knuckleprint and iris images.

References

[1] Aditya Nigam and Phalguni Gupta. A New Measure for Face Recognition System. In *5th International Conference on Image and Graphics (ICIG2009)*. IEEE CS, Xi'an, China, September, 2009.

[2] Aditya Nigam and Phalguni Gupta. Comparing Human Faces using Edge Weighted Dissimilarity Measure. In *11th International Conference on Control, Automation,Robotics and Vision (ICARCV 2010)*, Singapore December, 2010. included in IEEE Xplore.

Amrita Chaturvedi

Ontology is one of the most widely used semantic web technologies and is formally defined as "an explicit specification of a shared conceptualization". It can be used to model domain knowledge in the form of set of concepts and relationship between them and can be expressed in ontology languages like OWL, OIL, DAML etc.

Conceptual modeling is an activity undertaken during information systems development and maintenance work to build a representation of selected semantics about some real world domain. There are several domain modeling techniques in practice like Entity relationship modeling (ERM), Object Oriented Modeling (OOM), Object Role Modeling (ORM) etc. Recently ontologies have been vastly used for conceptual modeling purposes because of several benefits associated with them. However, we are still in need of elucidating links of differences between ontologies and conceptual schemas which might enable one to extend the existing techniques so as to obtain and represent maximum semantics from the concerned domain and/or to guide the developers in choosing suitable modeling technique.

Based on our study and analysis of existing conceptual modeling techniques i.e. ERM, OOM, ORM and conceptual modeling using ontologies, we differentiate between the existing CMT and ontologies based on several dimensions. This differentiation give us the leverage points on which we have built our discussion of advantages gained by ontological modelling with respect to the quality attributes of software systems. We, with the help of a case study have brought out the major differences between the relational and ontological model and have presented their quality attribute implications. This has enabled us to conclude the contexts/situations in which ontology should be used to model the domain knowledge (instead of traditional techniques) to enhance some particular quality attributes of software systems. The work is drafted in the form of a paper and will be sent in a related conference.

It is clear that use of ontologies in software architecture can change the quality attribute response of the architecture. We were more interested in finding out what change in quality attribute response can ontologies bring when used in software architectural styles. Some research has been done on role of ontologies in architectural styles like publish/subscribe system and semantic web services (SOA), so we thought of proceeding with the research on role of ontologies in Model-View-Controller architectural style.

We explored the different ways in which ontologies can be used into the MVC architectural style. The use of ontologies into the MVC architectural style gives rise to its new variant i.e. the 'semantic MVC' or 'Ontology Enabled MVC'. We also observe that a particular use of ontology into the MVC defines a new variant of observer pattern (which is the main ingredient of MVC style) i.e. the 'semantic observer pattern' or the 'ontology enabled observer pattern' which overcomes the drawbacks of observer pattern. We have also figured out the benefits gained by the use of ontologies in the MVC style.

Although we have identified the different ontologies and their uses in the MVC architectural style, there still remains the need to deeply investigate into the effect of these ontologies on the quality attributes of the style. This investigation of the effect of ontologies on the quality attribute of the MVC style is the work in progress.

Arpita Korwar

This year, I worked on the following topics:

1. Detecting bipartite matching in NC: The input is a bipartite graph with n vertices on each partition. The problem is to decide if a perfect matching exists in this graph or not. This problem can be decided in polynomial time (Edmonds '65). We are trying to find an NC algorithm that decides this problem.

2. Exact matching: This summer, I visited University of Ulm, Germany along with a colleague, Rohit Gurjar. We were visiting Dr. Thomas Thierauf and Dr. Jochen Messner. We worked on exact matchings. We are given a graph $G = (V, R, B)$, where V is the vertex set and the edges are colored either red or blue (the set of red edges is denoted by R and the set of blue edges is denoted by B), and an integer k. The problem is to decide if there is a perfect matching with exactly k red edges. We are studying various special cases of this problem, like planar graphs, complete bipartite graphs and multi-colored edges.

3. Presentation on "Operators in Computational Complexity" to the Calcutta Logic Circle: I presented the operators $\exists$ and $\forall$ over a complexity class C, through which, a hierarchy of complexity classes (for example, polynomial hierarchy) can be built. Probabilistic class BPP was defined, after which, we saw the $BP\cdot$ operator. I described interactive proof system (by Goldwasser-Micali-Rackoff) which can be used to build another hierarchy of complexity classes. A very similar hierarchy of complexity classes - the classes defined with Arthur-Merlin games (by Babai) was described using the new operators $\exists\cdot, \forall\cdot$ and $BP\cdot$. The complexity classes described using Arthur Merlin games collapse to a very small complexity class called AM. The proof for this was also presented.

4. Approximation algorithms workshop: I attended a workshop organized and sponsored by Microsoft Research. It included talks by illustrious people like Richard Karp, Sanjeev Arora and Naveen Garg. The talks were on topics like primal-dual method, LP rounding, PCP theorem, hierarchies of SDPs, etc.

5. Theory meet is an annual event where Indian researchers working in the field of Theoretical Computer Science meet and present talks on the current developments in the area. There were talks on Steiner graphs, parameterized complexity, communication complexity, maximum flows in undirected graphs, etc. I attended this meet.

Ashish Agrawal

During the year 2010-2011, I have completed my comprehensive exam and presented state of the art seminar on "Scalability in Cloud Computing". My research work is focused on software architecture and design patterns in softwares developed for cloud platforms. Cloud computing has given a new direction to software development by providing on-demand provisioning of hardware resources, virtualization and elasticity. However, existing software architectures, deployed on cloud platforms are not able to exploit its full power. Very few design patterns for cloud environment (like Map-Reduce) are introduced in the literature. Impact of cloud computing on the software architecture and development process is the problem domain.

We are working on the following problems, with the aim to enable software architects/programmers to design and write programs for cloud platforms:

1. What are the possible deployment architectures for a compute-intensive application running on cloud? We are working on providing grid as a service over the virtual machines of a cloud platform. Challenges are to improve performance and reduce overhead due to virtual machine isolation.

2. Identification of new design patterns (like Map-Reduce) for softwares developed for cloud platforms? We are currently working on few ideas taken from the hardware community; for example, speculative execution of tasks on another virtual machine to improve performance.

3. We are also comparing several cloud architectures for Software-As-A-Service. Candidate architectures differ with respect to deployment environment like VM-based virtualization, OS-based virutalization, combination of them, etc. for a 3-tier web application. Experiments are being done on IBM cloud machine and architectures are compared on several NFRs like scalability, performance, SLA etc.

4. Apart from cloud domain, I also finished my work in the area of business process management, which resulted in a declarative meta-model for describing business processes. The model is published as "Semantic of Business Process Vocabulary and Process Rules (SBPVR)" [1]. Processes modeled in SBPVR provides more flexibility and adaptability to the execution environment due to its declarative nature, rule-based approach and natural language representation.

Publications

[1] Ashish Agrawal. Semantics of business process vocabulary and process rules. In *4th India Software Engineering Conference (ISEC '11)*, Thiruvananthapuram, Kerala, India, 2011.

Badrinath G. S.

Progress of last academic year 2010-2011

During my last academic year 2010-2011, I have proposed some feature extraction techniques and indexing method for palmprint based recognition systems. It is been observed that the techniques proposed performs with accuracy of 100%. Further, the indexing technique proposed is robust to scale and rotation.

I have also written the first draft of my thesis which was corrected by my supervisor. Based on the comments, recently I have prepared the second draft and have given to the supervisor.

Following are the some of the publications obtained from the last academic year research work.

Publications

[1] Raghav Agrawal, G. S. Badrinath, and Phalguni Gupta. Image Enhancement Algorithm for Ink-on-paper Fingerprints. In *International Conference on Intelligent Computing*, Zhengzhou, China. August, 2011.

[2] Badrinath G. S. and Phalguni Gupta. Palm-print based Recognition System using Phase-Difference Information. *Journal of Future Generation Computer Systems*, 2011. Elsevier, In Press,.

[3] Badrinath G. S. and Phalguni Gupta. Some Efficient Feature Extraction Techniques for Palmprint. In *International Conference on Information Processing*, Bangalore, August, 2011.

Balwinder Sodhi

Research Status Report for Year 2010-11

Details

My research focus during the year 2010-11 has been around certain aspects of Cloud computing that concern software architecture and design. Since the past several years enterprises in various domains have shown increasing interest in moving towards network accessible services and applications. From social media and collaboration to SOA and data intensive back-end applications – all have seen emergence of interesting architecture and design paradigms. As a result of this, there has been many folds increase in the demands on system scalability, performance and agility, among other things. At the same time, the hardware and devices as well as the computing platforms have also seen a lot of innovation and emergence of new paradigms and technologies. Particularly, the virtualization and cloud oriented platforms have been quite disruptive technologies.

Given all of the above change and evolution, we wanted to explore the following:

1. How these changes in the modern datacenter and development environments have impacted the way we design and architect software systems and applications?

2. How can we characterize the computing environments to better understand their impact on software design and architecture activities?

3. How do various computing environments (e.g. virtual, cloud-oriented etc.) impact various non-functional quality attributes (NFQA) of a system?

4. How can we assess the suitability of a particular computing environment for a specific goal of satisfying certain NFQAs?

Results of our research investigations around the above questions have been published in few international conferences during this period [1, 2, 3].

Publications

[1] Balwinder Sodhi and T.V. Prabhakar. A design pattern to decouple data from markup. In *12th International Conference on Electronic Commerce and Web Technologies (EC-Web 2011)*, August 2011, Toulouse, France, 2011.

[2] Balwinder Sodhi and T.V. Prabhakar. Application architecture considerations for cloud platforms. In *Third International Conference on Communication Systems and Networks (COMSNETS 2011)*, January 2011, Bangalore, India, 2011.

[3] Balwinder Sodhi and T.V. Prabhakar. Assessing suitability of cloud oriented platforms for application development. In *9th Working IEEE/IFIP Conference on Software Architecture (WICSA 2011)*, June 2011, Boulder, Colorado, USA, 2011.

Deepanjan Kesh

Introduction

Let $k[x_1, \ldots, x_n]$ be a polynomial ring in n variables over the field k, and let $I \subset k[x_1, \ldots, x_n]$ be an ideal. Ideals are said to be *homogeneous*, if they have a basis consisting of homogeneous polynomials. *Binomials* in this ring are defined as polynomials with at most two terms [1]. Thus, a binomial is a polynomial of the form $c\mathbf{x}^\alpha + d\mathbf{x}^\beta$, where c, d are arbitrary coefficients. *Pure difference binomials* are special cases of binomials of the form $\mathbf{x}^\alpha - \mathbf{x}^\beta$. Ideals with a binomial basis are called *binomial ideals*. Toric ideals, the kernel of a specific kind of polynomial ring homomorphisms, are examples of pure difference binomial ideals.

Ideal Saturation

Problem Description

Saturation of an ideal, I, by a polynomial f, denoted by $I : f$, is defined as the ideal

$$I : f = \langle \{ \ g \in k[x_1, \ldots, x_n] \ : \ f \cdot g \in I \ \} \rangle.$$

Similarly, $I : f^\infty$ is defined as

$$I : f^\infty = \langle \{ \ g \in k[x_1, \ldots, x_n] \ : \ \exists a \in \mathbb{N}, f^a \cdot g \in I \ \} \rangle.$$

We describe a fast algorithm to compute the saturation, $I : (x_1 \cdots x_n)^\infty$, of a homogeneous binomial ideal I.

This problem finds applications in computing the radicals, minimal primes, cellular decompositions, etc., of a homogeneous binomial ideal, see [1]. This is also the key step in the computation of a toric ideal.

Our Approach

Before proceeding, we will need a few notations. U_i will denote the multiplicatively closed set $\{x_1^{a_1} \cdots x_{i-1}^{a_{i-1}} : a_j \geq 0, 1 \leq j < i\}$. $\prec_i$ will denote a graded reverse lexicographic term order with x_i being the least. $\varphi_i : k[x_1, \ldots, x_n] \to k[x_1, \ldots, x_n][U_{i-1}^{-1}]$ is the natural localization map $r \mapsto r/1$.

Algorithm 1 describes the saturation algorithm due to Sturmfels [2]. Algorithm 2 describes the proposed algorithm. The primary motivation for the new approach is that the time complexity of Gröbner basis is a strong function of the number of variables. In the proposed algorithm, a Gröbner basis is computed in the i-th iteration in i variables. The algorithm requires the computation of a Gröbner basis over the ring $k[x_1, \ldots, x_n][U_i^{-1}]$. The Gröbner basis over such a ring is not known in the literature. Thus, we propose a generalization of Gröbner basis, called Pseudo Gröbner basis, and appropriately modify the Buchberger's algorithm to compute it.

Data: A homogeneous binomial ideal, $I \subset k[\mathbf{x}]$.
Result: $I : (x_1, \ldots, x_n)^\infty$
1 **for** $i \leftarrow 1$ **to** n **do**
2 | $G \leftarrow$ Gröbner basis of I w.r.t. $\prec_i$;
3 | $I \leftarrow \langle \{ f : x_i^\infty \ : \ f \in G \} \rangle$;
4 **end**
5 **return** I ;
Algorithm 1: Sturmfels' Algorithm

Data: A homogeneous binomial ideal, $I \subset k[\mathbf{x}]$.
Result: $I : (x_1, \ldots, x_n)^\infty$
1 **for** $i \leftarrow 1$ **to** n **do**
2 | $G \leftarrow$ Pseudo Gröbner basis of $\varphi_i(I)$ w.r.t. $\prec_i$;
3 | $I \leftarrow \langle \{ \varphi_i^{-1}(f : x_i^\infty) \ : \ f \in G \} \rangle$;
4 **end**
5 **return** I ;
Algorithm 2: Proposed Algorithm

Radical and Minimal primes of an ideal

Radical of an ideal I, denoted by $\sqrt{I}$, is defined as the ideal

$$\sqrt{I} = \{ \ f \in k[x_1, \ldots, x_n] \ : \ f^m \in I, m \geq 0 \ \} .$$

A *prime* ideal is defined as an ideal I such that if $fg \in I$ implies either $f \in I$ or $g \in I$. The set of minimal primes of an ideal is the set of those primes containing I that are minimal w.r.t. inclusion.

Observation 1. *Let I be an ideal, and $\mathcal{P}$ be the set of minimal primes of I. Then*

$$\sqrt{I} = \bigcap_{P \in \mathcal{P}} P.$$

Problem Description

We describe a fast algorithm that computes the radical as well as the minimal primes of a binomial ideal.

Our Approach

Though the description of our solution is beyond the scope of this report, we can summarize some of the salient features of our solution -

Radical ideal Our solution has 2^n recursive calls compared to $n!$ in Eisenbud and Sturmfel's algorithm [1].

Minimal Prime Though our solution has to compute the same number of Gröbner basis as the algorithm due to Eisenbud and Sturmfels [1], the Gröbner basis in our case is computed in smaller rings than that of the former algorithm.

References

[1] D. Eisenbud and B. Sturmfels. Binomial Ideals. *Duke Mathematical Journal*, 84(1):1–45, 1996.

[2] B. Sturmfels. *Gröbner Bases and Convex Polytopes*, volume 8 of *University Lecture Series*. American Mathematical Society, December 1995.

Kamlesh Tiwari

Research Activities (July - 2010 to July 2011)

This year I started working in the field of biometrics and image processing. I had studied some of the effective feature extraction techniques and tried to develop familiarity in the area. We had also proposed a feature extraction technique for palmprint based identification system. In this system, the image is divided into disjoint sub-images. For each sub-image, the dominant orientation pixels based on the force field transformation are identified. Structure tensor values of these dominant orientation pixels of each sub-image are averaged to form tensor matrix for the sub-image. Eigen decomposition of each tensor matrix is used to generate the feature matrix which is used to take decision on matching. The system has been tested on IITK database. The experimental results reveal the accuracy of 100%. The result got published in International Conference on Intelligent Computing (ICIC-2011) [1].

References

[1] Kamlesh Tiwari, D.K. Arya, and P. Gupta. Palmprint based Recognition System using Local Structure Tensor and Force Field Transformation. In *International Conference on Intelligent Computing (ICIC)*, 2011.

Kiran Kumar Reddy

Automation Aspects of Architecture Decision Support System

From the transformation of use cases to UML views, some of the important open issues include providing automation support for:

1. Design alternative analysis.
2. Design decision analysis.
3. Transforming decision views to UML views.

Majority of our work includes proposing novel analysis models which are highly intuitive and automatable.

Pattern-oriented Knowledge Model for Architecture Design

Software design patterns document the most recommended solutions to recurring design problems. Selection of the best design pattern in a given context involves analysis of available alternatives, which is a knowledge-intensive task. Pattern knowledge overload (due to the large number of design patterns) makes such analysis difficult. A knowledge base to generate available alternatives can alleviate the problem. When architecture design knowledge is codified appropriately, the alternatives analysis problem can be modeled as an information retrieval problem. In this paper, we propose a pattern-oriented knowledge model which considers four dimensions of the pattern knowledge space: Pattern to Tactic relationship, Pattern to Pattern relationship, Pattern to Quality-attribute relationship and Pattern to Application-type relationship. We propose a graph based model to capture the semantics of a design pattern using design decisions and their dependencies. The relationships are analyzed using various graph properties which enable automation of relationship analysis. We perform analysis of these relationships for patterns in the two popular pattern catalogues viz GoF and POSA1.

A Game-theoretic Formulation of Tradeoff-points in Architecture Design

A software quality tradeoff is an important design context where multiple quality parameters are dependent on each other, a change in value in one quality parameter involves change in value in other quality parameters as well. At tradeoff points, designer analyzes various combinations of design alternatives and selects a combination which is considered as a balanced solution. Sometimes (e.g, under time pressure) designers simplify this analysis by selecting one quality parameter as most important and the design alternative combinations are assessed w.r.t. the most important quality parameter which in turn leads to biased design decisions. Biased design decisions can be sometimes undesirable because these can be one of the primary reasons for stakeholders' dissatisfaction. We propose a game-theoretic resource competition model to provide automation support for the designers following rationalistic decision making to achieve balanced decisions at tradeoff points during architecture design.

From Decision views to UML views - Refactoring approach

Decision view represents the design decisions and the dependencies among them. Decision view is umbrella view of different UML views. A design decision can be considered as a architecture transformation - additions, subtractions, modifications. A design decision dependency can be considered as transformation dependency, the subsequent design decision depends on the state change of precedent design decision. A design decision can affect one or more UML views. In order to determine which UML views a design decision effects, we use the quality attribute abstraction - The quality attributes of the UML views can be matched with the quality attributes affected by a design decision. Transformation on each of the views needs to be defined. We are planning to analyze the gaps in the UML views of various patterns.

Pawan Kumar Aurora

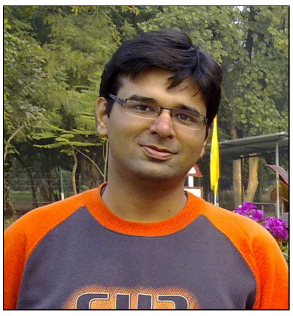
Research related activities for 2010-2011

I presented my State-of-the-art seminar on 'Coloring 3-colorable graphs' in September 2010.

We started working on the Minimum Graph Transformation problem. Our focus has been on finding an approximate solution using Semi-definite programming. Having designed the natural integer quadratic program for the problem we obtained the SDP relaxation. We got encouraging experimental results using the SDP and have been trying to theoretically justify the results. That has led us to studying the geometry of the feasible space as well as the geometry of the positive semi-definite cone. At this stage we have some conjectures that we hope on proving in the near future.

Of late we have started looking at the steiner network problem. We are focusing on the 2-approximation result due to Kamal Jain. His algorithm uses iterative rounding method that is required to solve the LP $O(m)$ times where m is the number of edges in the input graph. This is computationally prohibitive. There is no known combinatorial algorithm that achieves the same approximation guarantee. We plan to use the primal-dual schema to obtain a combinatorial algorithm for the problem. Another direction we can attempt is to reduce the number of iterations of Kamal Jain's algorithm in asymptotic terms.

I attended The 2011 School on Approximability organized by Microsoft Research India during January 05 - 09 at Bangalore.



Puneet Gupta

Annual Report for 2010-11-II

Courses done

During the first semester 2010-11-II did the following courses:

Course No.	Course Name
CS646	PARALLEL ALGORITHMS
EE608	DIGITAL VIDEO SIGNAL PROCESSING
CS797	SPECIAL TOPICS IN COMPUTER SCIENCE

Projects undertaken

I did the following project during my first semester 2010-11-II.

Calculating 2-D and 3-D images histogram and joint-histogram on CUDA architecture.
Calculating the image histogram of large-size 2-D and 3-D images is very time consuming and hence, applications involving these images are not of much utility in real time scenario. Hence, I used OpenMPI and CUDA for calculating the histogram of these images as fast as possible and analyses the bottleneck of using these.

Estimating the velocity of the vehicle in video image sequences.
Here,vehicle moving video database and it's corresponding camera calibration parameters are freely available. I used these parameters to estimate the vehicle velocity within a permissible limit of error upto 5%. Motivation behind this, is the automatic survillence.

3-D medical-images registration.
Here, the problem lies in the fact that we have different ways to capture images giving different features or modalities via different sensors, but, not a single sensor is capable of giving all desired features. This problem is faced in medical image processing. My work is to register these 3-D medical images so that valuable inference can be extracted by the experts of medical engineering domain.

Rohit Gurjar

Report of the work done in 2010-11

In year 2010-11 we mainly worked on bipartite matching and also looked on some other related problems. The sections below describes the progress in each of them.

Bipartite Matching

The aim here is to find an NC algorithm to detect a perfect matching in a bipartite graph. Our approach is to convert the problem into a special case of identity testing and then try to derandomize it. Our approach produced an alternate proof for an already known result that the problem is in NC if there are only polynomially many matchings in the graph. But till now, we have not been successful in this approach to get a general result.

Exact Matching

We first tried to get the same time complexity for Exact matching as Perfect matching for some particular versions and classes, but got no success there. However we could prove NC-equivalence between the following problems.

- (Bipartite) Perfect Matching
- (Bipartite) Maximum Matching
- Exact Matching in a complete (bipartite) graph

We are further investigating the exact complexity (NC^1 or NC^2 or NL) of these reductions.

Planar Matching

We tried to investigate if there is an NC reduction from Matching to Planar matching. We could show that a natural way (replacing each crossing in a planar embedding with a gadget) of doing this reduction is impossible.

Sagarmoy Dutta

Introduction

I am Sagarmoy Dutta, a Ph.D student in Computer Science and Engineering department of Indian Institute of Technology Kanpur. My research interests span computational algebra, complexity theory and quantum computing. Currently, under the supervision of my thesis guide Dr. Piyush P. Kurur, I am working on quantum error correcting code (QECC).

Overview of results obtained last year

We propose a new definition of cyclicity which applies to arbitrary quantum error correcting codes. This is directly motivated by the notion of cyclicity in classical linear codes. Recall that in classical scenario, cyclicity means invariance under the cyclic shift operation. In quantum setting, this shift operation on basis elements induces an unitary map which we call the quantum cyclic shift operator. We define a quantum code to be cyclic if it is invariant under quantum cyclic shift. We prove that in case of stabiliser codes our definition coincides with the familiar notion of cyclic stabiliser codes. Also, we show that the non-stabiliser $((n, 1+n(q-1), 2))_q$ codes given by Arvind. Kurur and Parthasarathy are cyclic according to our convention.

For explicit construction we focus on cyclic stabiliser codes over the field $\mathbb{F}_p$ whose lengths divide $p^t + 1$, for some positive integer t. We call such codes *t-Frobenius codes*, or just *Frobenius codes*, because of the key role played by the Frobenius automorphism. Restricting to lengths of similar form has previously appeared in coding theory. While constraining, it is not that bad, as there is a healthy, i.e. almost linear, density of such lengths. In bargain, it gives us a new way of tackling the symplectic inner product without using self-dual or Hermitian self-dual codes. We get a simpler formulation of the isotropy condition, which helps in the analysis of these codes considerably.

This simplicity of the isotropic condition also allows us to extend the notion of BCH distance for these codes and give efficient decoding algorithms. All our codes are uniquely cyclic and such codes cannot be CSS unless its distance is 1. Hence none of the code that we construct are CSS. Moreover, some of them are non-linear. This gives a family of codes for which efficient decoding algorithms were not known before.

We study the subfamily of *linear* Frobenius codes in detail and completely characterise them. This has two consequences, one negative and another positive. Firstly, over $\mathbb{F}_p$, we show that there are no t-Frobenius linear codes when t is odd. This is a somewhat serious limitation of linear cyclic codes as the density of such lengths n seems to be almost linear. Moreover, this impossibility is purely Galois theoretic unlike other known restriction that arise from sphere packing bounds or linear programming bounds.

On the positive side, the characterisation of linear Frobenius codes gives us ways to explicitly construct examples of linear Frobenius codes of lengths $p^{2t} + 1$. Again, since the density of such lengths are also healthy, this technique gives sizable number of explicit examples including the well studied Laflamme code.

Since there cannot be any linear t-Frobenius code for odd t, it is natural to ask whether we can construct nonlinear codes of such length. Using a generalisation of our technique for liner Frobenius codes, we show that it is indeed possible. Linear stabiliser codes are based on classical codes over quadratic extension of the base field. Our generalisation crucially makes use of higher degree extensions which results in explicit examples of nonlinear t-Frobenius codes for both odd end even t.

Publication

The results stated above are accepted as a paper named *Quantum Cyclic Code of Length Dividing* $p^t + 1$ and co-authored by Piyush P. Kurur in the *IEEE International Symposium on Information Theory* (ISIT 2011). It is to appear in the conference proceeding which will will be hosted in IEEE xplore.

Travel support by Research I foundation

I visited Max Planck Institute, Saarbrucken, Genmany as a guest scientist from 27/9/2010 to 16/12/2010. The travel expenses and dearness allowance of 30 days were born by Research I foundation. Also I attended ISIT 2011 which was held in St. Petersburg, Russia from 31/7/2011 to 5/8/2011. This visit was entirely funded by Research I foundation.

Satyam Sharma

I am a third year Ph.D. student working in the area of cryptology and security. My other areas of interest are computational complexity, concurrent data structures and algorithms, operating systems and computer networks.

In the past year, I have continued to work in the area of secure multi-party computation. Secure multi-party computation deals with the problem of efficiently computing an n-ary randomized functionality that maps n local inputs to n local outputs in a setting consisting of n different parties that communicate by passing messages while preserving certain security properties such as independence of inputs, privacy, output delivery, correctness and fairness. Theoretical general constructions that achieve the above goals and security properties under suitable assumptions have long been well known, but the challenge now is to make them practically useful both as generic tools as well as for specific functionalities. Some opportunities I had identified after my state-of-the-art seminar last year were:

1. Can we improve the aforementioned general protocols by optimizing their local computation and interaction requirements?

2. Can we construct specialized protocols (that are simpler and more efficient than the general ones) that deal with specific practically useful functionalities?

We are exploring the first question by aiming to improve the above general protocols by improving one of the key tools used by them, verifiable secret sharing. The idea of computing over encrypted data is closely related (and equivalent) to secure multi-party computation. Recently there have been advances in this area with the invention of a fully homomorphic encryption scheme (the sum or product of the encrypted values is the encrypted sum or product of the individual values, respectively) that makes it possible to compute arbitrary functions of inputs that are encrypted without decrypting the initial or intermediate values. Analogously, a fully homomorphic secret sharing scheme (the sum or product of the shares of the secrets is the share of the sum or product of the secrets, respectively). Such a fully homomorphic secret sharing scheme would preclude the need for interaction and the transfer of intermediate secret shares after each step during secure multi-party computation. We are exploring whether such secret sharing schemes are possible, and if they exist, how can they be constructed?

The progress we have made so far is to come up with novel secret sharing schemes that are only additively or multiplicatively homomorphic (many other such schemes have been known for long). The challenge remains to construct a secret sharing scheme that is simultaneously additively and multiplicatively homomorphic.

Saurabh Joshi

I have published a paper in International Conference on Distributed Computing and Networking (ICDCN) 2011[1] held at Bangalore in Jan 2011, abstract of which is appended below.

Abstract

Barrier synchronization is widely used in shared-memory parallel programs to synchronize between phases of data-parallel algorithms. With proliferation of many-core processors, barrier synchronization has been adapted for higher level language abstractions in new languages such as X10 wherein the processes participating in barrier synchronization are not known a priori, and the processes in distinct "places" don't share memory. Thus, the challenge here is to not only achieve barrier synchronization in a distributed setting without any centralized controller, but also to deal with dynamic nature of such a synchronization as processes are free to join and drop out at any synchronization phase. In this paper, we describe a solution for the generalized distributed barrier synchronization wherein processes can dynamically join or drop out of barrier synchronization; that is, participating processes are not known a priori. Using the policy of permitting a process to join only in the beginning of each phase, we arrive at a solution that ensures (i) Progress: a process executing phase k will enter phase k + 1 unless it wants to drop out of synchronization (assuming the phase execution of the processes terminate), and (ii) Starvation Freedom: a new process that wants to join a phase synchronization group that has already started, does so in a finite number of phases. The correctness of the solution is formally established. From the perspective of a global observer, our protocol guarantees a bound of at most two phases from the phase a process had registered its intention to join. We show how the testing by each of the processes with all the other processes can be short circuited leading to efficient synchronization. The above protocol is further generalized to multiple groups of processes (possibly non-disjoint) engaged in barrier synchronization.

I went to Microsoft Research Lab, India from May 2010 to Sept 2010 for research collaboration. We worked on automatic inferrence of atomic sections of concurrent programs. Earlier work has focussed on inferring such regions with the help of user annotations. In few other works where user annotations are not required, there is no guarantee of soundness and/or completeness. We propose a sound and complete method of inferring such atomic sections without needing annotations from the user. The same work will be submitted to a suitable venue soon.

References

[1] Shivali Agarwal, Saurabh Joshi, and Rudrapatna K. Shyamasundar. Distributed generalized dynamic barrier synchronization. In *Proceedings of the 12th international conference on Distributed computing and networking*, ICDCN'11, pages 143–154, Berlin, Heidelberg, 2011. Springer-Verlag.

M Seetha Ramaiah

Academic activities during Aug 2010- Aug 2011

Recently I have changed my research area from Software Architecture (specifically Autopoietic Computing) to Machine Learning and its applications. During the last 8-9 months I have been familiarizing myself with several Machine Learning methods and related literature.

Since May 2011, I have been reading literature related to deep learning [1]. One of the goals of deep learning is to come up with a hierarchy of representations of the input data where the lower level representations of the hierarchy are composed to form the higher level representations so that the highest level will be a very low dimensional representation of the original input. This idea in particular is inspired by the way human brain works. For example, the visual cortex does not process the sensory data at one shot; instead it passes them through a series of layers which learn some lower dimensional representations of the input.

Other motivations for deep learning come from complexity theory. A particularly interesting result from complexity theory is by Håstad et al. [3], which effectively says the following: certain functions which can be represented by a depth k network with polynomially many nodes, as a function of the input size, require exponentially many nodes when we try to represent them by a depth k-1 network. A network with an exponential number of nodes implies learning an exponential number of parameters. That requires a substantially larger training set. Learning a deep network with traditional back-propagation based approaches requires a huge amount of training data because they are discriminative methods and try to get all the information about the input data from the corresponding labels, which usually don't give as much information. Hinton et al., [2] give a greedy layer-wise algorithm that first trains a series of Restricted Boltzmann Machines in an unsupervised manner, one layer at a time, and then fine tune the weights using a traditional back-propagation method. The unsupervised pre-training gives a good starting point for the back-propagation algorithm to search for global minima, and that leads to quicker convergence and also a better overall performance.

References

[1] Yoshua Bengio. Learning Deep Architectures for AI. In *Foundations and Trends in Machine Learning*, volume 2(1), pages 1–127. Now Publishers, 2009.

[2] G. E. Hinton, S. Osindero, and Y Teh. A fast learning algorithm for deep belief nets. *Neural Computation*, 18:1527–1554, 2006.

[3] Johan Håstad and Mikael Goldmann. On The Power Of Small-Depth Threshold Circuits. *Computational Complexity*, 1:610–618, 1991.

Shubhadip Mitra

Research Progress Report 2010-2011

Semester I (July-Dec., 2010)

In the first semester, I registered for three courses namely

1. CS628 COMPUTER SYSTEMS SECURITY,
2. CS687 ALGORITHMIC INFORMATION THEORY, and
3. CS697 SPECIAL TOPICS IN COMPUTER SCIENCE with Prof. Arnab Bhattacharya

A brief description of the CS697 research work [1] is as following:

Problem determination is essentially diagnosis of root cause of a problem in a system. As systems are evolving over the years, they are becoming increasingly large, dynamic and distributed in nature. Problem Determination in such complex systems is a challenging task. In this work, we investigate this area and propose a framework for problem determination which is based on detection of complex events over event streams generated by different components in the distributed system. A problem may be represented by a pattern graph where the nodes represent the event attributes such as timestamps, type, location; and the edges represent the event relationships like temporal distance, spatial distance or causal. Our problem reduces to searching for a problem pattern over event streams. Any implementation of this framework requires efficient algorithms for the constraint satisfaction problem which is NP complete. A few efficient algorithms have been proposed for this problem under different settings namely offline, online and distributed. The algorithms for tree patterns perform very well. Using these as building blocks, we suggest a hierarchical heuristic for general problem patterns. This work is an extension of our prior work [3] and is still under progress.

A poster on the above work was presented at IBM ICARE 2010 conference held at Bangalore.

I finished this semester with an SPI of 8.67.

Semester II (Dec-April, 2011)

In this semester, I registered for CS 797 with Prof. S. K. Mehta and three units of thesis with Prof. Arnab Bhattacharya. I spent time in picking up the domain knowledge required for my PhD research. A short description of the CS797 research work [2] is as following:

There are applications which seek a path in a graph that satisfy multiple constraints. In this work, we studied two such problems, namely the bounded path weight problem and the bounded path bottleneck problem on directed acyclic graphs with fixed dimension vector weights on its edges. The bounded path weight problem seeks a path between a fixed pair of nodes, whose weight must satisfy certain lower and upper bounds specified in each component. We show that the problem is NP complete and propose a fully polynomial time approximation scheme(FPTAS) to solve it. We also discuss two applications: bounded non-linear objective path problem and bounded subgraph homeomorphism problem. The bottleneck value

of a path is the value of the minimum weighted edge on the path. We extend this definition to define the bottleneck vector of a path when the graph has vector weights on its edges. The bounded path bottleneck problem determines a path between a specified pair of vertices when some lower and upper bounds are imposed on its bottleneck values in each component. We present a polynomial time algorithm to solve this problem on directed acyclic graphs. We are working towards publishing this work in a suitable journal.

I secured an SPI of 10 in this semester.

References

[1] Shubhadip Mitra. Event Pattern based Framework for Problem Determination in Distributed Systems. *Tech. Report*, 2010.

[2] Shubhadip Mitra. Multi-Constrained Path Problems on Directed Acyclic Graphs with Applications. *Tech. Report*, 2011.

[3] Shubhadip Mitra, Partha Dutta, Shivkumar Kalyanaraman, and Prashant Pradhan. Spatio-Temporal Patterns for Problem Determination in IT Services. *Services Computing, IEEE International Conference on*, 0:49–56, 2009.

Sudhanshu Shukla

Report of the work done in 2010-11

During the year 2010-2011, I successfully completed my comprehensive exam and presented state of the art seminar on "Future of Microprocessors : Challenges and Opportunities". The main focus of my research has been scalability of multi-core processors. Multi-core processors with upto sixty four cores are already commercially available and future microprocessors are expected to have hundreds/thousands of cores. The scalability of multi-core processors is generally constrained by off-chip memory bandwidth. Better techniques for on-chip memory management techniques are required to fully exploit the computational throughput provided by future thousand core microprocessors.

A multi-core processor system behaves much like a multi-node system, but exhibits important differences. With increased level of on-die integration of cores in multi-core processor, the storage available on-die is much more restricted, while the communication latencies are considerably lower. The reduced communication latencies have led to major changes in two aspects of multi-core processor systems, when compared to a large scale multi node system. First, thread migration between computing cores has become less costly and can be done entirely in hardware today. Second, on-die data migration between computing core has become less costly. This can also be done entirely in hardware without any intervention from the operating system. These properties of multi-core processor systems raise the question that whether traditional directory-based cache coherence protocols present the optimal design for future multi-core processors.

Since January 2011, I have been working under the supervision of Dr. Mainak Chaudhuri, trying to address the above question of optimal cache coherence protocol for future multi-core processors. We have been working on designing cache coherence protocol which exploits both thread migration and data replication.

Sujith Thomas

First semester 2010-11

Comprehensive Oral Examination

During my first semester I gave my **Comprehensive Exam**. The title of my presentation was *'String Similarity Search'*. The talk covered the area of string similarity search using *positional q-grams*. We also discussed the construction of B^{ed}-tree, a B^{+}-tree based index structure for evaluating all types of similarity queries on edit distance and normalized edit distance between strings.

Thesis

On the thesis front I explored the various psychological theories surrounding concepts and their related issues. The literature I reviewed include the works of Laurence, Margolis, Rosch etc. I looked at the probabilistic models for word learning proposed by Tenenbaum et. al. Also, I read literature on Bayesian Networks.

Second semester 2010-11

State of the Art Seminar

During my second semester I gave a seminar titled **'Representing Concepts'**. I discussed the following knowledge representation techniques in my talk - Formal Concept Analysis, Conceptual Graphs and Probabilistic Models of Cognition. I also discussed the advantages and disadvantages of each representation based on the psychological properties that a knowledge representation should satisfy.

Thesis

I reviewed the theory of conceptual spaces proposed by Gärdenfors. I read the 'Probabilistic Model of Theory Formation' proposed by Kemp et. al. I gave a talk on 'Multiagent Inductive Learning' where agents perform inductive learning based on the process of argumentation. I also explored the area of Description Logic and probabilistic Description Logic.

Surya Prakash

A Rotation and Scale Invariant Technique for Ear Detection in 3D[1]

Introduction

Ear based human recognition has received much attention in recent years because the nature of the ear is stable and it provides a reliable way for human recognition [5]. It has been found that ear is invariant to different facial expressions and is unaffected by aging unlike face. It is also unaffected by cosmetics and eye glasses which is the case for face and iris respectively. Ear background is also predictable as it always remains fixed at the middle of the profile face. Moreover, ear can be captured easily without the cooperation of the subject. It can be employed in an stand alone human recognition system or can be integrated with the face for enhanced recognition. In spite of ear having so many rich features compared to other biometrics, reported accuracy for 2D ear recognition has firmly kept it away from being widely used. However, the use of 3D ears have helped in improving the accuracy of ear recognition and recently many systems have been proposed with good recognition accuracy [3, 8, 6, 4].

Ear recognition in 3D consists of two major phases namely (i) Ear detection from the profile face and (ii) Recognition. Most of the well known ear recognition techniques proposed in the literature (*viz.* [2, 7, 6]) have focussed on recognition phase and have used manually cropped ears for recognition. There exist a few techniques to detect and crop ear automatically in 3D profile face range images. However, most of these techniques need a registered 2D ear image for ear detection [3, 8, 4]. Also, these techniques are not very efficient in ear detection, particularly when profile face images are affected by scaling and rotation (pose variations). Moreover, they are not fully automatic to be deployed in realtime scenarios. For an efficient ear recognition system, specially in non-intrusive applications, it is very much required to automatically detect and crop the ear from a whole profile face image which may be affected by scale and pose variations.

Detection of ears from an arbitrary 3D profile face range image is a challenging problem because ear images can vary in scale and pose (due to in-plane and out-of-plane rotations) under various viewing conditions. In this research work, we attempt to develop a technique which can overcome these issues and can provide an efficient scale and rotation invariant solution for automatic ear detection in 3D profile face range images.

Technique Overview

The technique is based on the fact that in a 3D profile face range image, ear is the only region which contains maximum depth discontinuities; as a result, this place is rich in edges. It also relies on the fact that edges belonging to an ear are curved in nature. The technique makes use of graph connected components constructed using the edge map of the profile face range image for ear detection. Main focus of this work is to develop a technique with following characteristics:

1. It should be able to perform ear detection in 3D without using a registered 2D image.

[1] This work is in progress and final results are expected soon.

2. It should be invariant to rotation (in-plane and out-of-plane) and scale.

3. It should detect left and right ear simultaneously without imposing any additional training cost.

Preliminary Findings

To analyze the performance of the technique, initial experiments are conducted on University of Notre Dame public database, Collection J2 (UND-J2) [1] which consists 3D profile face range images with scale and pose variations. Preliminary ear detection results obtained by the technique are encouraging. We are process of carrying out a detailed experimentation.

Publication Details

Paper Title : An Enhanced Geometric Hashing
Authors : Umarani Jayaraman, Amit Kumar Gupta, Surya Prakash and Phalguni Gupta
Conference : IEEE International Conference on Communications (ICC 2011), Kyoto, Japan, June 5-9, 2011.

Abstract

This paper presents an enhanced geometric hashing technique suitable for object recognition. Unlike the available geometric hashing, the proposed technique needs less amount of time and memory, uniform index distribution in the hash space without using any rehashing function. It performs indexing and searching in one pass with linear complexity. The proposed technique has been applied in biometric databases. It has been tested for three traits such as ear, iris and palm print. The hit-rate of 100% has been achieved for top 5 best matches in all cases.

References

[1] University of Notre Dame Profile Face Database, Collection J2. http://www.nd.edu/ cvrl/CVRL/DataSets.html

[2] Bhanu, B., Chen, H.: Human ear recognition in 3D. In: Proc. of Workshop on Multimodal User Authentication, pp. 91–98 (2003)

[3] Chen, H., Bhanu, B.: Human ear recognition in 3D. IEEE PAMI **29**(4), 718–737 (2007)

[4] Islam, S.M.S., Davies, R., Bennamoun, M., S. Mian, A.: Efficient detection and recognition of 3D ears. IJCV (2011). , DOI:10.1007/s11263-011-0436-0

[5] Jain, A.K., Flynn, P., Ross, A.A. (eds.): Handbook of Biometrics, chap. 7. Springer-Verlag New York, Inc. (2007)

[6] Passalis, G., Kakadiaris, I.A., Theoharis, T., Toderici, G., Papaioannou, T.: Towards fast 3D ear recognition for real-life biometric applications. In: Proc. of IEEE Conference on Advanced Video and Signal Based Surveillance (AVSS' 07), pp. 39–44 (2007)

[7] Yan, P., Bowyer, K.: Ear biometrics using 2D and 3D images. In: Proc. of Int'l Conference on Computer Vision and Pattern Recognition-Workshops, pp. 121–128 (2005)

[8] Yan, P., Bowyer, K.: Biometric recognition using 3D ear shape. IEEE PAMI **29**(8), 1297–1308 (2007)

Umarani Jayaraman

Efficient search and retrieval techniques in multi-modal biometric database

Multimodal biometrics recognition system uses multiple sources of biometric information to establish the identity. The system is reliable and robust due to the presence of multiple, fairly independent pieces of evidence. In the identification mode, the system recognizes an individual by searching the templates of all the users in the database for a match of user biometric data. The trivial solution for identification problem is to search all images in the database against a query image. The process to fetch each image from the database and to compare it against the query image for a match is computationally inefficient. Hence, there is a need of indexing of database images so that the search space and the complexity of identification can be reduced. But, there are some critical issues which should be handled while indexing images of a large database. Some of them are mentioned below.

1. Generally, there are too many features in an image.
2. In multimodal biometrics recognition system the methodology adopted to integrate the information from multiple biometric traits.
3. To reduce the time complexity to process and match two images.

To reduce the time complexity to process and match two images the design adopted was to index the images based on their score vectors (i.e, similarity score). The idea is simple, in case of large biometric databases, matching between two images is much slower than computing the correlation coefficient between two score vectors. Since the match score vectors also high in dimension VA+ file has been used to index the score vectors. The above work resulted in a paper [3] which gives a detailed explanation of the design.

An efficient indexing scheme has been proposed that can be used for retrieval from a large iris database. For a given color iris query image, the proposed indexing scheme makes use of iris color to determine an index and uses this index to reduce the search space in the large iris database. Further, for query q, the retrieval technique uses iris texture to find the top best match from the reduced search space. This work resulted in a paper [1] which gives a detailed explanation of the design.

An enhanced geometric hashing technique has been proposed which is suitable for subject recognition. Unlike the available geometric hashing, the proposed technique needs less amount of time and memory, uniform index distribution in the hash space without using any rehashing function. It performs indexing and searching in one pass with linear complexity. This work resulted in a paper [2] which gives a detailed explanation of the design.

Published Papers

[1] Umarani J, Surya Prakash, and Phalguni Gupta. An Iris Retrieval Technique Based on Color and Texture. In *Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP 10)*, Chennai, India, December, 2010.

[2] Umarani J, Surya Prakash, and Phalguni Gupta. An Enhanced Geometric Hashing. In *IEEE Communication and Information Systems Security Symposium (ICC 2011)*, Kyoto, Japan, February, 2011.

[3] Ashish Paliwal, Umarani J, and Phalguni Gupta. A Score based Indexing Scheme for Palmprint Databases. In *17th IEEE International Conference on Image Processing (ICIP 10)*, Hong Kong, September, 2010.