

Jornada de Iniciação Científica e Tecnológica do LNCC

Petrópolis, 25 de setembro de 2014.

Laboratório Nacional de Computação Científica – LNCC

Diretor

Pedro leite da Silva Dias

Coordenação de Administração - CAD

Anmily Paula dos Santos Martins

Coordenação de Ciência da Computação - CCC

Jauvane Cavalcante de Oliveira

Coordenação de Matemática Aplicada - CMA

Sandra Mara Cardoso Malta

Coordenação de Mecânica Computacional - CMC

Regina Célia Cerqueira de Almeida

Coordenação de Sistemas e Controle - CSC

Gilberto de Oliveira Corrêa

Coordenação de Sistemas e Redes – CSR

Wagner Vieira Léo

Programa Institucional de Bolsas de Iniciação Científica &
Programa Institucional de Bolsas de Iniciação em Desenvolvimento
Tecnológico e Inovação

Renato Portugal

Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq

Presidente

Glaucius Oliva

Coordenadora Geral do PIBIC

Lucimar Batista de Almeida

Jornada de Iniciação Científica e Tecnológica do LNCC

Comissão Interna do PIBIC/LNCC

Renato Portugal
Eduardo Lucio Mendes Garcia
Helio José Correia Barbosa
Jack Baczynski

Avaliador Externo

Demerson Nunes Gonçalves
UCP

Apresentação

Apresentamos neste volume a relação dos trabalhos da Jornada de Iniciação Científica e Tecnológica do LNCC, desenvolvidos pelos bolsistas no período de agosto de 2013 a julho de 2014. Estes trabalhos foram apresentados em sessão pôster para, juntamente com as fichas de avaliação do bolsista, serem utilizados como elementos de avaliação do PIBIC/PIBITI-LNCC. Nesta avaliação contamos com o Prof. Dr. Demerson Nunes Gonçalves - UCP, que em função das entrevistas e exame deste volume preparou e enviou ao CNPq relatório de avaliação do Programa.

Agradecemos o empenho dos Professores Orientadores e dos Bolsistas na preparação das fichas de avaliação e dos pôsteres e da disponibilidade da presença para entrevista, para que nosso programa fosse avaliado. É preciso mencionar que temos sido bem avaliados, estando nosso PIBIC/PIBITI bem colocado entre os melhores do país, conforme divulgação feita pelo CNPq.

Aproveitamos a ocasião para agradecer ao CNPq pelas bolsas concedidas, a Direção do LNCC pelo apoio e a Comissão Interna deste programa no LNCC.

Renato Portugal
Coordenador do PIBIC/PIBITI - LNCC

Índice

Implementação de rotinas paralelas para o ambiente SPiNMe-MHM.....	1
Bolsista: <u>Abel Pinto Coelho</u>	
Orientador: Antonio Tadeu	
Ambiente de cluster para gerência de dados científicos.....	3
Bolsista: <u>Adolfo Simões Neto</u>	
Orientador: Fabio Porto	
Estudo, teste e implementação de bibliotecas para visualização 3D.....	5
Bolsista: <u>Alexandro de Paula Miranda</u>	
Orientador: Jauvane C. de Oliveira	
Gerador WEB de Workflows Científicos - Flowman.....	7
Bolsista: <u>Christian Meirelles dos Santos Mury</u>	
Orientador: Bruno Schulze	
Avaliação de desempenho dos algoritmos para manipulação de grafos em ambientes paralelos híbridos CPU/GPU.....	8
Bolsista: <u>Daniel Nascimento Ramos da Silva</u>	
Orientador: Carla Osthoff	
Um modelo epidemiológico com sazonalidade e imunidade cruzada.....	13
Bolsista: <u>Dérek Prates</u>	
Orientador: Maurício Kritz	
Coorientador: Jaqueline Maria Silva	
Content-Based Image and Video Retrieval by Statistical Pattern Recognition.....	15
Bolsista: <u>Daniel Nascimento</u>	
Orientador: Carla Osthoff	
Estudo da implementação de computação quântica.....	24
Bolsista: <u>Edgar Winter da Costa</u>	
Orientador: Renato Portugal	
Uso de estimativas de precipitação por radar a bordo de satélite na Região Serrana do Rio de Janeiro.....	25
Bolsista: <u>Fabio Magest Filgueiras</u>	
Orientador: Pedro Leite da Silva Dias	
Implementação do campo de força MMFF94 no programa de predição de estrutura de proteínas GAPF.....	29
Bolsista: <u>Frederico José Rodrigues Carlos</u>	
Orientador: Laurent Dardenne	
Computação Quântica.....	30
Bolsista: Hugo Souza	
Orientador: Renato Portugal	

Introdução À Computação Quântica e Geração de Números Aleatórios.....	34
Bolsista: <u>Igor de Souza Costa</u>	
Orientador: Renato Portugal	
Construção de um sistema de controle para o conjunto cyberforce.....	40
Bolsista: <u>Johnathan Fercher da Rosa</u>	
Orientador: Jauvane de Oliveira	
Reconhecimento de padrões e interpolação de dados com aplicações para imagens médicas.....	46
Bolsista: <u>Larissa Alves Pereira</u>	
Orientador: Gilson Giraldi	
Estudo de desempenho de uma aplicação de bioinformática em um ambiente híbrido de processamento paralelo com tecnologica multicore e GPU.....	49
Bolsista: <u>Mateus Silva Melo</u>	
Orientador: Carla Osthoff	
Coorientador: Roberto Souto	
Controle e simulação de sistemas sujeitos a saltos markovianos.....	50
Bosista: Oscar Neiva Neto	
Orientador: Marcos Todorov	
Avaliação de desempenho do paralelismo de um problema estreitamente ligado em um ambiente multicore Intel com e sem Hyper-Threading.....	53
Bolsista: <u>Rafael Lourenço</u>	
Orientador: Carla Osthoff	
Exploração de ambientes virtuais com navegação Natural.....	54
Bolsista: <u>Raquel Ellem Marcelino de Oliveira</u>	
Orientador: Jauvane Cavalcante	
Visualização estereoscópica em monitores desktop.....	57
Bolsista: Richard Daniel Calile Salles	
Orientador: Jauvane Cavalcante	
Modelagem Geométrica e Otimização de objetos 3D para ambientes virtuais colaborativos.....	59
Bolsista: <u>Stephane Deserto Vasconcelos</u>	
Orientador: Jauvane Cavalcante	
Modelagem da Manipulação do Comportamento do Hospedeiro – Caso Toxoplasma Gondii.....	62
Bolsista: <u>Tainara Mendes de Andrade Soares</u>	
Orientador: Maurício Vieira Kritz	
MHOLline 2.0 – A evolução de um workflow científico para problemas em Bioinformática e Biologia Estrutural.....	64
Bolsista: <u>Victor Crisóstomo Cruz Reis</u>	
Orientador: Laurent Dardenne	
Coorientador: Priscila Capriles	

Modelagem Computacional Aplicada a Questões Sócio-Educacionais.....68

Bolsista: **Weslei Peter de Oliveira**

Orientador: José Karam Filho

Coorientadores: Fabiano Saldanha Gomes de Oliveira



Implementação de rotinas paralelas para o ambiente SPiNMe-MHM

Bolsista: Abel Pinto Coelho de Souza

Orientador: Antonio Tadeu Azevedo Gomes (LNCC/MCTI)

O trabalho desenvolvido nos sete primeiros meses (janeiro-julho de 2014) de atuação como bolsista de iniciação científica (IC) consistiu na prototipação de uma abordagem que facilite a escalabilidade do trabalho gerado por uma classe de aplicac , aumentando a eficiência das suas soluções.

O desenvolvimento de novos métodos numéricos é de grande importância nas Ciências Computacionais, Matemáticas e da Engenharia. Infelizmente, a implementação dos métodos mais utilizados na atualidade consome muito tempo enquanto gostaria-se de aprofundar a exploração nas propriedades numéricas do método (acurácia, por exemplo) e/ou fazer comparações com outros métodos. No ambiente computacional SPiNMe (Scientific Productivity in Numerical Methods) [1], a especificação, implementação e experimentação dos métodos podem ser facilitadas, comparando-se com o trabalho que dedica-se hoje para essas atividades na prática usual de desenvolvedores de código para métodos numéricos, prática esta que envolve fundamentalmente o emprego de linguagens de programação estáticas como Fortran e C/C++ (fortemente tipadas e executáveis por meio de compilação prévia para o ambiente de execução pretendido) e bibliotecas cuja extensibilidade é restrita em uma mais dimensões de definição dos métodos numéricos (discretização do domínio, formulação variacional, funções de base etc.)

No SPiNMe os usuários podem implementar os métodos numéricos desenvolvidos, testá-los e, dentro do ambiente, comparar seus resultados com métodos previamente implementados no próprio ambiente. Como estratégia para esse fim emprega-se no SPiNMe uma linguagem dinâmica cuja interpretação, embutida em uma biblioteca de elementos finitos típica, permite a exploração das definições do método numérico através de mudanças, em tempo de execução, do código interpretado associado a essas definições.

Em função do exposto acima, partiu-se para a implementação de uma versão alternativa do ambiente SPiNMe, especificamente desenhada para dar suporte a família de métodos MHM. Essa versão



do ambiente (SPiNMe-MHM) desacopla a atividade de distribuição de subproblemas entre processos (intensiva de comunicação), que é implementada usando a linguagem Erlang, da atividade de cálculo numérico efetuado sobre cada subproblema (intensiva de computação), que é implementada usando a linguagem C++. Nesse novo desenho arquitetural, o paralelismo no nível inter-processo ocorre de forma desacoplada entre os processos responsáveis pelo cálculo numérico sobre cada subproblema em um mesmo nível (escala). É importante contudo destacar que, ainda que a estratégia de “dividir para conquistar” viabilize a geração de subproblemas de tal forma que a complexidade computacional dos processos responsáveis pelo cálculo numérico sobre esses problemas seja compatível com a capacidade computacional dos processadores que hospedarão cada processo, a implementação desses processos deve ser tal que a capacidade computacional dos processadores disponíveis (hoje quase que exclusivamente de natureza multiprocessada) seja efetivamente utilizada.

Com o trabalho proposto, foi possível automatizar o processo de compilação do ambiente SPiNMe-MHM, bem como testar uma implementação específica do método MHM. É possível mostrar que há espaço considerável para maiores ganhos de desempenho se compararmos com a execução do mesmo método sem a utilização da camada de distribuição (implementada em Erlang) do ambiente SPiNME-MHM.



Ambiente de cluster para gerência de dados científicos

Bolsista: Adolfo Simões

Orientador: Fábio Porto

O termo em inglês *Big Data* refere-se à grandes volumes de dados requerendo novas técnicas para o tratamento de seu armazenamento, processamento e análise [5,1]. No domínio de aplicações científicas, *Big Data* se traduz no tratamento dos dados produzidos durante o ciclo de vida da pesquisa, desde a formulação da hipótese científica até sua confirmação ou refutação. No estágio atual da gerência de dados científicos, têm-se focado em uma das etapas deste ciclo, referente a fase do ciclo experimental [2]. Nesta fase, os cientistas modelam seus experimentos utilizando-se de linguagens de *workflows* científicos [7,2] e, com muita frequência, mantêm os dados em arquivos do sistema operacional. O processamento desses arquivos é realizado por sistemas de gerência de *workflows* científicos, tais como Kepler [3], Swift [6], Chiron [3] e QEF [4].

Este plano de trabalho teve por objetivo a instalação e configuração de um *cluster* para aplicações intensivas de dados, visando atender problemas de natureza científica utilizando o paradigma *MapReduce*. Dentre as atividades desenvolvidas pelo bolsista destacam-se: (i) amplo estudo de instalação, testes e levantamento de requisitos do *cluster*; (ii) instalação da rede básica, criação de *host* virtual e configuração de *bridges*; (iii) criação e configuração dos nós *masters* e nós *slaves*; (iv) instalação do *Hadoop*, *Postgresql*, *Lustre*, *HBASE*, *CentOS*, *Apache server* e, (v) carga de dados da astronomia nos SGBD's para teste.

Termo esta experiência de Iniciação Científica muito satisfeito, pois adquiri conhecimentos muito importantes para o meu desenvolvimento profissional/científico. Abrindo novas e importantes oportunidades, e contribuindo para minha formação na área de Big Data em linha com os desafios do laboratório para os próximos anos.

- Referências bibliográficas

[1] BECLA J., LIM K-T. Report from The First Workshop on Extremely Large Databases, Stanford Linear Accelerator Center, Menlo Park, CA, USA, 2007.



- [2] MATTOSO, M., WERNER, C., TRAVASSOS, G. H., et al. Towards Supporting the Life Cycle of Large Scale Scientific Experiments. *International Journal of Business Process Integration and Management*, v. 5, pp. 79-92, 2010.
- [3] OGASAWARA, E., Daniel de Oliveira, Patrick Valduriez, Jonas Dias, Fabio Porto, Marta Mattoso: An Algebraic Approach for Data-Centric Scientific Workflows. *PVLDB* 4(12): 1328-1339 (2011).
- [4] PORTO F., TAJMOUATI O., SILVA V. F. V., SCHULZE B., AYRES F. V. M. QEF - supporting complex query applications, 7th IEEE International Symposium on Cluster Computing and the Grid — CCGrid 2007 , Rio de Janeiro, Brazil, pp. 846-851, 2007.
- [5] STONEBREAKER M., BECLA J., DEWIIT D. et al. Requirements for Science Data Bases and SciDB. *CIDR Perspectives*, 2009.
- [6] WILDE, M., HATEGAN, M., WOZNIAK, J., et al., Swift: A language for distributed parallel scripting, *Paralell Computing*, v. 37, n. 9, pp. 634-652, 2011.
- [7] ZHAO, Y., RAICU, I., FOSTER, I., Scientific Workflow Systems for 21st Century, New Bottle or New Wine?" In: *Proceedings of the 2008 IEEE Congress on Service Part I*, pp.467- 471, Honolulu, Jul. 2008.



Estudo, teste e implementação de Bibliotecas para visualização 3D

Laboratório Nacional de Computação Científica (LNCC)

Alexsandro de Paula Miranda ¹, Jauvane C. de Oliveira²

1. Introdução

Ambientes Virtuais Colaborativos (AVC) tem sido, historicamente, aplicados a diversas áreas do conhecimento, tais como simulação militar de combate para treinamento de tropas, design e engenharia, treinamento, engenharia de software, medicina, etc. Tais aplicações, se imersivas, podem incluir um nível a mais de realismo em uma simulação, permitindo que o usuário tenha uma experiência similar àquela que o mesmo teria em similar situação no mundo real. Aplicações de grande escala, que dependem da transferência, em tempo real, de uma massa de dados considerável para visualização colaborativa pelos participantes da simulação impõem uma carga considerável no subsistema de comunicação, o que as torna inviáveis para uso com redes convencionais.

2. Atividade realizada pelo bolsista

O ambiente CAVE (Cave Automatic Virtual Environment) é uma interface de realidade virtual, formada por paredes, teto e piso nas quais podem ser visualizadas imagens 3D projetadas (CRUZ-NEIRA, 1992), e permite ao usuário a sensação de estar de fato imerso no mundo sintético, porém essa sensação de imersão está diretamente relacionada à coerência das imagens apresentadas em três dimensões. São usadas bibliotecas de visualização 3D que não possuem uma compatibilidade satisfatória com o ambiente CAVE. Para isso uma pesquisa foi realizada para encontrar bibliotecas existentes que poderiam ser modificadas para atender as necessidades do AVC do LNCC.

3. O ambiente CAVE do LNCC

O Laboratório Nacional de Computação Científica (LNCC) que se encontra na cidade de Petrópolis, no estado do Rio de Janeiro, possui um ambiente CAVE em uma de suas instalações. Atualmente a ferramenta de desenvolvimento do sistema de controle para o ambiente CAVE é o Instant Reality, arquitetura que combina vários recursos, fornecendo uma interface para desenvolvedores de aplicações de Realidade Virtual e Aumentada (REALITY, 2010). O objetivo desta pesquisa é encontrar bibliotecas de visualização 3D que possam servir como uma segunda via para o ambiente CAVE que só possui o Instant Reality como sistema de controle.



3.1 Escolhendo, testando e implementando

Nas primeiras implementações realizadas utilizou-se o OGRE 3D (Object-oriented Graphics Rendering Engine), um motor gráfico open source escrito em C++ que permite a manipulação de objetos 3D. Inicialmente havia uma motivação para o uso do Ogre, devido a compatibilidade com bibliotecas de simulação de física que é um recurso desejável para uso em aplicações do ambiente CAVE. No entanto há um grau mais elevado na curva de aprendizado e há pouco material bibliográfico disponível, o que é um aspecto negativo para o uso desta ferramenta.

Começou-se então uma nova implementação utilizando a arquitetura Vizard que também é um motor gráfico escrito em Python, não é open source, e possui uma imensa biblioteca disponível a fim de facilitar a implementação. Essas bibliotecas influenciaram na evolução dos projetos que outrora eram mais complexos de se desenvolver.

4. Os resultados

Sabe-se que as lâmpadas dos projetores do ambiente CAVE possuem tempo de vida útil. Para evitar um consumo desnecessário dessas lâmpadas utilizou-se dois monitores de 20 polegadas para iniciar as implementações, pois assim haveria um melhor aproveitamento das telas quando todas as viewports fossem geradas, simulando as paredes do CAVE. Com o ambiente computacional preparado devidamente, distribuiu-se todas as viewports pela tela com tamanho equivalente as paredes do CAVE, assim ao criar uma viewport com X de altura e Y de largura na parede do CAVE, no monitor o seu tamanho seria devidamente proporcional e vice-versa. Configurou-se também a manipulação dessas viewports pelo teclado para que fosse possível obter o tamanho exato em tempo real, assim elas não ultrapassariam as arestas do ambiente CAVE. Feito isso, foi necessário alinhar o objeto dentro do ambiente CAVE, de modo que, quando um objeto fosse movido ele passaria de uma viewport para outra, dando uma sensação de continuidade.

5. Projetos Futuros

Continuar pesquisando a respeito de RV para que seja possível unir as várias aplicações existentes do laboratório ACiMA (Ambientes Colaborativo e Multimídia Aplicada) em uma única aplicação no Vizard, utilizando sockets.



Gerador WEB de Workflows Científicos - Flowman

Bolsista: Christian Meirelles dos Santos Mury
Orientador: Bruno Schulze e Matheus Bandini

Um sistema de *workflow científico* é uma forma especializada de um *workflow*, composto por processos e dados, que são manipulados obedecendo uma sequência pré-definida, com o objetivo de auxiliar a solução de um problema científico. Os *workflows científicos* se caracterizam e diferem dos demais tipos de *workflows*, por sua complexidade e normalmente pela grande quantidade de dados processados em cada uma de suas etapas.

O objetivo do trabalho proposto foi analisar o processamento de *workflows* e desenvolver um sistema capaz de gerar um esquema estruturado, para uso no processamento de *workflows científicos*. Para isso é disponibilizado aos seus usuários, uma interface gráfica acessada por meio de um navegador, onde ele é capaz de determinar o sequenciamento das etapas do *workflow científico* (processos e dados), sendo possível editá-la, arquivá-la e recuperá-la, por meio da criação e manipulação de um diagrama de blocos.

Com isso foi possível disponibilizar uma interface que auxilia seus usuários na criação e visualização das etapas que constituem um *workflow científico*, gerando o esquema necessário ao seu processamento, e permitindo por meio do uso da tecnologia de *containers* a criação de ambientes virtuais pré-configurados e customizados com as aplicações necessárias a execução do *workflow científico* gerado pelo usuário.

Este trabalho teve como resultado prático um protótipo capaz de permitir aos usuários manipularem as aplicações necessárias a execução de um *workflow* na forma de blocos, que têm associados estas aplicações prontas para serem executadas em ambientes virtualizados contidos em *containers*, contribuindo assim para o esforço da prospecção do uso da computação em nuvem em prol da computação massivamente paralela e distribuída.



Avaliação de Desempenho dos Algoritmos para Manipulação de Grafos em Ambientes Paralelos Híbridos CPU/GPU.

Daniel Nascimento Ramos
da Silva

dramos@lncc.br

Carla Osthoff

osthoff@lncc.br

Frederico Luís Cabral

fcabral@lncc.br

Hoje em dia, uma das vertentes da computação que tem ganhado cada vez mais proeminência é a de aplicação da computação massivamente paralela, isto é, o uso de aceleradores gráficos - GPUs (*general processing units*) – no processamento de computação em geral. O crescimento que se nota nesse tipo de computação deve-se principalmente ao fato de que a arquitetura de processadores gráficos é desenhada para realização intensiva de cálculos aritméticos. As GPUs podem conter até milhares de ULAs (unidades lógicas aritméticas) – arquitetura *manycore* - se encaixando em problemas complexos aptos à paralelização (KIRK e HWU, 2013).

No mundo contemporâneo, problemas envolvendo grafos – estrutura de dados usada geralmente para representação de relacionamentos de rede – surgem frequentemente (CORMEN, LEISERSON, *et al.*, 2009). Geralmente, aplicações que envolvem tal tipo de estrutura demandam tempos de execução elevados, o que vai de encontro ao uso da computação paralela, visando soluções a prazos mais curtos.

Esse trabalho destina-se a resolução computacional de problemas que envolvem o uso de grafos utilizando uma abordagem em GPU e, à avaliação de desempenho comparativa das soluções desenvolvidas com soluções em ambientes seriais e paralelos mais tradicionais como arquiteturas de memória distribuída, memória compartilhada e híbrida entre as duas últimas (MATLOFF, 2014).

Como ferramenta de desenvolvimento tem sido utilizada a plataforma de computação paralela de propósito geral CUDA™ (*Compute Unified Device Architecture*) (NVIDIA® CUDA™, 2014). CUDA™ é uma solução tanto de *software* quanto de *hardware* que é mantida e desenvolvida pela Nvidia® para sua linha de aceleradores gráficos.

Os problemas tratados nesse trabalho são: a) árvores geradoras de amplitudes mínima; b) caminhos mais curtos de única origem; c) caminhos mais curtos de todos os pares; d) fluxo máximo; e) busca em largura; e) conectividade ST (BONDY e MURTY, 2008). Além dos problemas supracitados, esse trabalho implementa uma solução em GPU do cálculo local de centralidade de redes complexas



(CLCRC) como caso de estudo.

A respeito do andamento desse trabalho:

- i) Tem sido efetuada pesquisa e aprofundamento na tecnologia CUDA, teoria dos grafos e teoria das redes complexas;
- ii) Como passo de adaptação à tecnologia, foram implementados os algoritmos de busca em largura e conectividade ST propostos em (HARISH, VINEET e NARAYANAN, 2009), sendo que enquanto o primeiro já foi testado e está funcional enquanto o segundo encontra em fase de resolução de problemas concernentes a sua implementação;
- iii) O foco principal nesse trabalho tem sido a implementação de uma solução em GPU do CLCRC.

A metodologia base para a solução do CLCRC é a DACCER (*Distributed Assessment of the Closeness Centrality*) proposta por (WEHMUTH e ZIVIANI, 2013), implementada em CUDA por (MOREIRA, 2013) e testada analítica e comparativamente em alguns dos ambientes de computação paralela em (CABRAL, OSTHOFF, *et al.*, 2014) como parte desse trabalho.

A metodologia DACCER avalia a centralidade de um nó em uma rede complexa, através de uma vizinhança estabelecida por um raio h de distância. Por exemplo, quando $h=0$ são considerados no cálculo do volume (medida adotada pelo DACCER) os vizinhos ligados diretamente ao nó em questão, já quando $h=1$ são considerados no cálculo os vizinhos ligados por até duas arestas ao nó. Nessa metodologia foi identificado que $h=2$ demonstra um valor de correlação satisfatório entre a classificação gerada pelo DACCER e metodologias globais (que utilizam métricas de conhecimento total da topologia da rede) (SCARDONI e LAUDANNA, 2012) (WEHMUTH e ZIVIANI, 2013).

Entretanto, apesar de que para $h=0$ e $h=1$ a solução CUDA é executada em alguns casos até mais de 100 vezes mais rápido do que soluções seriais e paralelas, a implementação do DACCER tem tido sérios problemas quando o $h=2$: divergência de fluxo de execução e estouro de memória no ambiente do *device* (GPU).

A fim da adequação do DACCER à plataforma Nvidia®, foram implementadas relaxações no método. A principal relaxação concerne a permissividade de vizinhos repetidos a cada nó. Por exemplo, se um vizinho de um nó aparece na vizinhança de $h=0$ e $h=1$, na metodologia DACCER, para o cálculo de volume do nó, esse vizinho era considerado apenas uma vez. Na nova metodologia denominada MDACCER (*Modified DACCER*) esse vizinho é considerado as duas vezes. Esse tipo de abordagem,



fez com que diversas estruturas de dados e *kernels* (funções executadas no *device*) que mantinham e controlavam as vizinhanças fossem inutilizadas, assim como, a divergência de fluxo vista na remoção de vizinhos repetidos também foi eliminada com essa estratégia.

Como houve mudanças no algoritmo, foram efetuados testes de correlação entre MDACCER e DACCER e, entre MDACCER e as estratégias de CLCRC globais *Closeness Centrality* e *Betweenness Centrality*. Os resultados de correlação em modelos de Redes Barabási-Albert são em muitos casos superiores a 90% utilizando a estratégia *Ranking Biased Overlap* (WEBBER, MOFFAT e ZOBEL, 2010).

Em relação aos testes de desempenho foram utilizadas 4 soluções: a) DACCER serial; b) DACCER em MPI utilizando 32 processos; c) DACCER em GPU; d) MDACCER em GPU. As soluções a) e b) foram implementadas em (WEHMUTH e ZIVIANI, 2013), enquanto a solução c) em (MOREIRA, 2013). O teste foi executado no cluster Altix-XE (LNCC, 2014) do LNCC tendo o h variado de 0 a 2 e sendo utilizadas modelos de Redes Barabási-Albert.

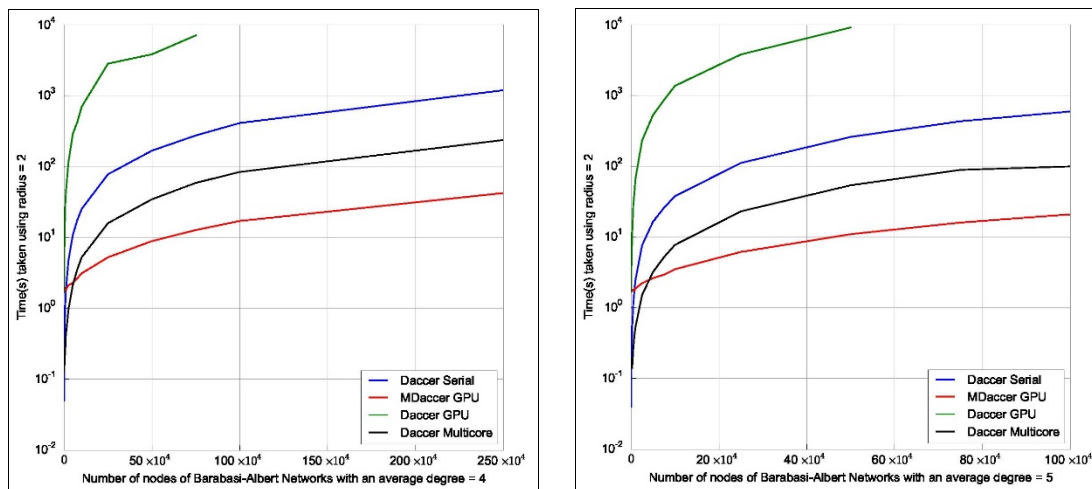
Foi verificado nos testes que a metodologia MDACCER utilizando $h=2$, conforme o tamanho das redes crescia, é executada em tempos inferiores as demais soluções.

A Figura 1 ilustra tal cenário, demonstrando o ganho de desempenho dela em relações as demais soluções: a) contra redes com grau médio = 4 e b) contra redes com grau médio = 5. Para alguns tamanhos de rede nota-se um *speedup* de mais de 4x contra a solução em MPI e mais de 20x contra a solução serial

Figura 1 - Desempenho comparativo para $h=2$

a)

b)



Fonte: Autoria própria

Como próximo passo desse trabalho, desenvolver-se-á uma solução genérica da metodologia MDACCER, visando a análise de redes maiores, mais densas e raios maiores (redes com números elevados de vértices e/ou arestas, e raios maiores estouram a memória disponível na GPU). Essa solução genérica, em suma, visa particionar o problema de acordo com a memória disponível no dispositivo, assim é apta a analisar uma gama maior de redes complexas e ser executada em placas gráficas com menor memória disponível.

Paralelamente a esse passo, far-se-á testes com traços de redes reais e outros modelos de rede como o proposto por (ERDOS e RÉNYI, 1959). Conseqüente a esses passos continuar-se-á a implementação de problemas envolvendo grafos, primeiramente a continuação ao algoritmo conectividade ST e posteriormente o problema de caminhos mais curtos de única origem.

Referências Bibliográficas

ALBERT, R.; BARÁBASI, A.-L. Statistical mechanics of complex networks. **Reviews of Modern Physics**, v. 74, n. 1, p. 47-97, jan. 2002.

BONDY, J. A.; MURTY, U. S. R. **Graduate Texts in Mathematics Series: Graph Theory**. New York: Springer, 2008.



- CABRAL, F. L. et al. Massive Parallelism with GPUs for Centrality Ranking in Complex Networks. **International Journal of Computer Science & Information Technology (IJCSIT)**, Tamil Nadu, 6, n. 3, jun. 2014. Disponível em: <<http://airccse.org/journal/jcsit/6314ijcsit02.pdf>>.
- CORMEN, T. H. et al. **Introduction to Algorithms**. 3ª. ed. Cambridge: MIT Press, 2009.
- ERDOS, P.; RÉNYI, A. On random graphs I. **Publicationes Mathematicae**, Debracen, 6, 1959. p. 290–297.
- HAGBERG, A.; SCHULT, D.; SWART, P. Exploring Network Structure, Dynamics, and Function using NetworkX. **Proceedings of the 7th Python in Science Conference (SciPy2008)**, Pasadena, CA USA, Agosto 2008. 11–15.
- HARISH, P.; VINEET, V.; NARAYANAN, P. J. **Large Graph Algorithms for Massively Multithreaded Architectures**. International Institute of Information Technology. Hyderabad. 2009.
- HOBEROCK, JARED; BELL, NATHAN. Thrust: A Parallel Template Library - Version 1.7.0, 2010. Disponível em: <<http://thrust.github.io/>>.
- KIRK, D. B.; HWU, W.-M. W. **Programming Massively Parallel Processors: A Hands-on Approach**. 2ª. ed. Waltham: Elsevier, 2013.
- LNCC. Cluster Altix-XE, 2014. Disponível em: <<http://www.lncc.br/altix-xe/>>. Acesso em: 03 Setembro 2014.
- MATLOFF, N. **Programming on Parallel Machines**. Los Angeles: University of California, 2014. Disponível em: <http://heather.cs.ucdavis.edu/matloff/public_html/158/PLN/ParProcBook.pdf>.
- MOREIRA, R. N. **Paralelismo Massivo no Cálculo da Centralidade dos nós de Redes Complexas com o uso de GPUs**. Teresópolis: Trabalho de Conclusão de Curso (Graduação) - Ciência da Computação - Centro de Ciência e Tecnologia Centro Universitário Serra dos Órgãos, 2013.
- NVIDIA® CUDA™. CUDA Parallel Computing Platform. **NVIDIA**, 2014. Disponível em: <http://www.nvidia.com/object/cuda_home_new.html>. Acesso em: 3 jul. 2014.
- SCARDONI, G.; LAUDANNA, C. Centralities Based Analysis of Complex Networks. **New Frontiers in Graph Theory**, Rijeka, mar. 2012. 324-348. Disponível em: <<http://www.interchopen.com/books/new-frontiers-in-graph-theory/centralities-based-analysis-of-networks>>.
- WEBBER, W.; MOFFAT, A.; ZOBEL, J. A similarity measure for indefinite rankings. **ACM Transactions on Information Systems**, Amsterdam, 28, Novembro 2010. 1-38.
- WEHMUTH, K.; ZIVIANI, A. Distributed Assessment of the Closeness Centrality. **Computer Networks**, Philadelphia, 57, set. 2013. 2536-2548.



Um modelo epidemiológico com sazonalidade e imunidade cruzada

Prates, D.B.¹, SILVA, J. M.¹, Kritz, M. V.²

1- Instituto de Ciência, Engenharia e Tecnologia, ICET, UFVJM.

2- Laboratório Nacional de Computação Científica, LNCC.

RESUMO

Em muitos países, especialmente na América do Sul e Ásia, a Dengue é um problema de saúde pública, uma vez que apresenta altas taxas de infecção e mortalidade, principalmente em épocas de chuva, além de não existir uma vacina ou medicamento de proteção contra a doença. A Organização Mundial da Saúde (OMS) estima que entre 50 e 100 milhões de pessoas são infectadas anualmente pelo vírus da Dengue em mais de 100 países em todos os continentes, exceto a Europa.

A Dengue apresenta a peculiaridade de ser causada por até quatro tipos virais. Assim, um indivíduo infectado por um tipo viral apresenta imunidade total a ele e uma imunidade parcial e temporária para os outros três. Além disso, a Dengue tem uma característica sazonal forte, porque o seu vector encontra condições favoráveis de reprodução distintas ao longo do ano.

A modelagem matemática oferece recursos para uma análise qualitativa de uma epidemia. Os modelos mais eficientes são baseados em equações diferenciais e propõem simular a dinâmica de uma população infectada com uma doença, permitindo a inserção de parâmetros que permitem representar características específicas da epidemia estudada, tais como, taxas de infecção ou de mortalidade associadas.

O modelo SIR, que foi o pioneiro na modelagem epidemiologia, foi proposto por Kermack e McKendrick em 1927, que propõe dividir a população em três classes: suscetíveis, infectados e removidos (ou recuperados). Com a inserção de indivíduos infectados na população, a epidemia pode ser analisada apresentando resultados conclusivos. O modelo é constantemente modificado para simular situações mais realistas e biológicas.

Para o estudo epidemiológico da Dengue novos parâmetros e classes foram inseridos no modelo SIR original, permitindo que o novo sistema de equações diferenciais modele a situação epidemiológica, abrangendo os seguintes conceitos: a co-circulação simultânea dos tipos virais da Dengue; a imunidade cruzada, que é a imunidade parcial de um variante viral para que o indivíduo que foi previamente infectado com uma outro variante viral; e a sazonalidade, que representa a tendência de uma epidemia de reincidência ou intensificar-se em um determinado período de tempo.

É observado que tais mudanças representam bem a situação epidemiológica relacionada à Dengue. O objetivo deste trabalho é a utilização de tal modelo modificado, inserção da sazonalidade, para analisar a epidemia da Dengue, que tem condições climáticas favoráveis para a propagação da doença.

Referências



- [1]. Aguiar, M., Stollenwerk, N., Kooi, B.W., 2009. Torus bifurcations, isolas and chaotic attractors in a simple dengue model with ADE and temporary cross immunity. *Int. J. Comput. Math.* 86, 1867–1877.



Content-Based Image and Video Retrieval by Statistical Pattern Recognition

D. N. M. Cardoso, Jacques Facon and G. A. Giraldi

Abstract – With the Internet evolution, there has been a huge increase in the amount of images and videos stored in electronic format in the World Wide Web. Nowadays, large image and video databases are available which poses fundamental requirements for data retrieval. Such requirement can be implemented using pattern recognition techniques. In this paper, we consider the problems of image and video retrieval by content-based techniques. Firstly, we describe a framework for image retrieval that considers an SVM ensemble based on the "one-against-all" SVM multi-class approach. The feature space is generated by the discrete cosine transform. Given a query image, the SVM approach performs its classification and the Euclidean distance is used to compute the similarity between the query and the images in the same class. Next, a video retrieval pipeline is described. It follows a strategy based on object segmentation, key frame as well as video shot detection, and feature extraction. In the video segmentation step, we consider region growing and quadtree techniques.

Keywords – CBIR, SVM, SBIM, Web system.

Introdução

In the last decades, the spread of high quality acquisition hardware and internet facilities for data accessibility and storage have motivated the generation of large image and video database which are available in the World Wide Web. As a consequence, techniques have been developed for manipulating, retrieving and sharing images and videos over the Internet. In this paper we focus on content-based image and video retrieval.

This field had a fast development in the past decades and involves algorithms for processing, feature extraction, classification and similarity measure [4]. It is an important tool for users that need to search for information inside an image or video database. In this paper we firstly review the work presented in [1] by some of us for image retrieval. We propose a solution based on Support Vector Machine (SVM) ensemble built using the "one-against-all" SVM multi-class approach described in [10]. In the feature extraction step we get a compact representation of the image by using the Discrete Cosine Transform (DCT) of the image. The DCT implementation is simpler than other image transforms (like the wavelets) and we have obtained suitable results with DCT. Then, following [10], we construct N SVM models, one for each class of the database.

Given the query image, this SVM ensemble is used to find the class for the query classification. Specifically, each SVM returns a real number that is normalized and interpreted as the probability of the

Este trabalho foi desenvolvido no Laboratório Nacional de Computação Científica (LNCC), Petrópolis, RJ, em colaboração com o Laboratório de Visão Computacional do PPGIA da Pontifícia Universidade Católica do Paraná, campus de Curitiba.

D. N. M. Cardoso: graduado em Análise de Sistemas pela UFPR e atualmente aluno de graduação no curso de ciência da computação pela Pontifícia Universidade Católica do Paraná. (douglasmeireles@gmail.com).

G. A. Giraldi: Pesquisador do Laboratório Nacional de Computação Científica (LNCC), Petrópolis, Brasil (gilson@lncc.br).

Jacques Facon: Professor Adjunto da Pontifícia Universidade Católica do Paraná (PUCPR), na escola politécnica (facon@ppgia.pucpr.br).

query to belong to the corresponding class. So, we select the class with the largest probability. This class will be used in the final step for image similarity computation and retrieval. Specifically, the Euclidean distances from the query image to the class images is used as a similarity measure for image retrieval.

The video retrieval technique follows the methodology presented in [7]: (a) the input video stream is segmented;

(b) Key frame construction based on the visual contents; (c) Detection of fundamental sequences in the video stream (shot detection); (d) Feature extraction; (e) Similarity measure to compare the query video clip with the database videos. In [7] the input raw video is segmented using the Expectation Maximization (EM) algorithm in order to extract the objects presented in the raw video. Then, the obtained result is processed for shot detection and to extract feature vectors using texture analysis and color moments. Finally, all the extracted features are combined as a feature vector. This process is applied for all the database videos as well as for the query video clip in order to allow a comparison between the query and the database videos. This comparison is performed via the feature weighted distance measure and the similar videos are retrieved from the database.

We are in charge to implement the step (a) using an alternative technique based on quadtree [8] and region growing [3]. The paper is organized as follows. The section 2 performs the background review. In sections 3 and 4 we describe the content-based image and retrieval methodologies, respectively. Next, we present preliminary results using the proposed video segmentation method. Finally, in section 5, we offer the final comments.

Background

In this presentation a true-color RGB image I is represented by a (generalized) matrix $I \in \mathbb{R}^{M \times N \times 3}$, which means that it has M rows and N columns and depth equal to 3. Such a matrix can be represented by a column vector $x \in \mathbb{R}^{M \cdot N \cdot 3}$ obtained by concatenating its elements column-by-column.

When applying an orthogonal image transform $C \in \mathbb{R}^{M \times N}$ to an image $I \in \mathbb{R}^{M \times N \times 3}$ we apply the transformation to each channel separately; that means: $CIC^T = (CI(\cdot, 1)C^T, CI(\cdot, 2)C^T, CI(\cdot, 3)C^T)$ [5]. Given a discrete transform $\in \mathbb{R}^{M \times N \times 3}$, the zonal mask defines a filtering operation given by:

$$\tilde{V}(k, l) = M(k, l) \cdot V(k, l), \quad (1)$$

where the matrix M , the mask, is given by: $M(k, l) = 1$ if $1 \leq k, l \leq R$ and $M(k, l) = 0$, otherwise; with R been an application-dependent parameter.

Among the possibilities for orthogonal (or unitary) transforms [5], the DCT is (in general) efficient for compression using a zonal mask to select the highest image transform coefficients (in absolute values). The obtained result can be represented by a column vector $x \in \mathbb{R}^{R \cdot R \cdot 3}$ and used as the input for the SVM classifier.

The seeded region growing performs a segmentation of an image by checking the neighboring pixels of specific points, known as seed points, in order to determine whether the pixels could be classified to the cluster of the seed point or not [3].

A quad tree is a spatial data structure in which each internal node has exactly four children (see Figure 1). The root of the tree corresponds to the entire image which is divided into four quadrants. Each quadrant is analyzed using a pre-defined criterion in order to decide if it will be divided also. The process proceeds until no further subdivisions are possible [8].

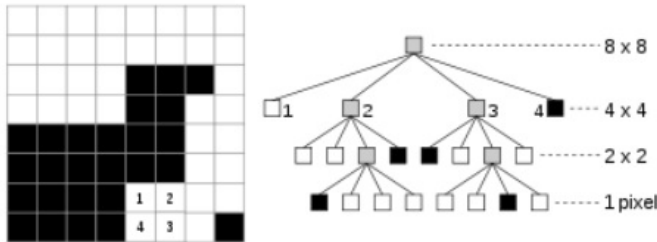


Figure 1 Quadtree scheme (Source: <http://en.wikipedia.org/wiki/Quadtree>).

A linear SVM [9] is primarily a two-class classifier that maximizes the width of the margin between classes, that is, the empty area around the separating hyperplane defined by the distance to the nearest training samples. It can be extended to multi-class problems by solving essentially several two-class problems.

Given a training set that consists of L pairs of $(x_1, y_1); (x_2, y_2) \dots (x_L, y_L)$, where x_i denotes the n -dimensional training observations and $y_i \in \{-1, 1\}$ the corresponding classification labels, the SVM method [9] seeks to find the hyperplane defined by:

$$f(x) = (x \cdot w) + b = 0, \quad (2)$$

Algorithm 1 SVM Ensemble Construction

Input: Image Classes set: $\Phi = \{C_1, C_2, \dots, C_N\}$; SVM parameters.

for $i = 1 \rightarrow N$. do

 Take all k images from class i and label them as 1.

 Random sampling $(2k)/(N - 1)$ images from the other classes and label them as -1 .

 Generate training set S of expression (5).

 Compute SVM machine i by expression (4).

end for

Output: SVM machines $SV M_1, SV M_2, \dots, SV M_N$.

which separates positive and negative observations with the maximum margin. It can be shown that the solution vector

$$w_{\text{svm}} = \sum_{i=1}^L \alpha_i y_i x_i, \quad (3)$$

where α_i are non-negative coefficients obtained by solving a quadratic optimization problem with linear inequality constraints [9]. Those training observations x_i with non-zero α_i lie on the boundary of the margin and are called support vectors.

In order to generalize the SVM we must consider separating hyper surfaces for classification. The

idea is to map the original input vectors x into a high-dimensional *feature space* Z through some nonlinear mapping such that in the Z space an optimal separating hyperplane can be constructed.

This idea can be implemented by using a kernel $K(x, y)$ that generates the inner product for the feature space. In this case, we can generalize expression (2) by using the inner product defined by the kernel K [9]:

$$f(x, a) = \text{sign}(\sum_{\text{support vectors}} y_i \alpha_i^0 K(x_i * x) + b) \quad (4)$$

Content-Based Image Retrieval

In this section we summarize the methodology for CBIR described in [10] which has the following stages: (a)

Feature extraction; (b) SVM ensemble construction; (c) Classification of the query image; (d) Image retrieval.

In the feature extraction process we firstly compute the DCT for each color band. As usual for DCT image transform [5], we observe that highest image transform coefficients can be selected using a zonal mask given by the expression (1). The parameter R is experimentally defined and depends on the image patterns.

After the feature extraction process, each image of the database is represented by a column vector $x_i \in \mathbb{R}^{M \times N \times 3}$. So, let N be the total number of classes. Like in [10], we firstly construct N support vector machines using these feature vectors through a “one-against-all” SVM multi-class approach [10].

Specifically, for training each SVM model i , we take all k images from class i and label them as 1. Then, using random sampling we choose $(2k)/(N-1)$ images from classes other than i and label them as -1 [10]. The obtained set of feature vectors and corresponding labels l_m :

$$S = \{(l_1, x_1), (l_2, x_2), \dots\}, \quad x_m \in \mathbb{R}^n, \quad l_m \in \{-1, 1\}. \quad (5)$$

with $n = R \cdot R \cdot 3$, is the training set for SVM model i , denoted by SVM_i . The algorithm (1) summarizes the steps for generating all the N SVM models.

After generating the SVM models $SVM_1, SVM_2, \dots, SVM_N$, given a query image z , we can find the estimated labels $y_i, i = 1, 2, \dots, N$, using these machines:

$$\begin{aligned} y_1 &= z_1 w_1^1 + z_2 w_1^2 + \dots + z_n w_1^n, \\ y_2 &= z_1 w_2^1 + z_2 w_2^2 + \dots + z_n w_2^n, \\ &\dots \\ y_N &= z_1 w_N^1 + z_2 w_N^2 + \dots + z_n w_N^n. \end{aligned}$$

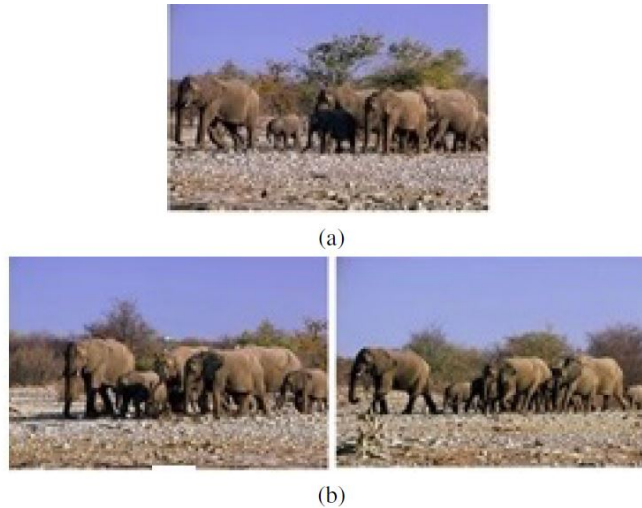


Figure 2 - (a) Query image. (b) Most similar images returned by the retrieval engine (Image source: COREL Database).

Then, we adopt the strategy of normalizing each label y_j and interpret it as the probability that the query image z belongs to class j . Then, we set the class label for image z by using the highest normalized estimate $y_i^j, i = 1, 2, \dots, N$.

Then, like in [10], it is calculated and sorted the Euclidian distances between the query image and all the images that belong to the same class. Images with the lowest Euclidean distances are considered similar images and returned by the system. This completes the image retrieval step of our method.

The work [1] presents results obtained with this methodology as well as its integration in the Shared Biological Image Manager (SBIM) system. This system allows users to have great control over the images of its databases, giving them the possibility to group images by category, search images by size, format, compression and color channel. It also has a message manager, to allow the communication among researchers [6].

The Figure 2 pictures a result obtained with the implemented methodology. The Figure 2 (a) shows a sample of the COREL database ([alipr.com/cgi-in/zwang/regionsearch show.cgi](http://alipr.com/cgi-in/zwang/regionsearch_show.cgi)) and the Figures 2.(b) present the most similar samples found by the proposed method.

Content-based video retrieval

The methodology for video retrieval to be implemented is presented in [7] and can be summarized

in the Figure 3. The basic steps are also presented in the image retrieval scheme presented on section 3; that is: feature extraction, classification of the query image, and data retrieval. However, video retrieval imposes some specific challenges inherent to the video stream processing. For instance, videos have much

richer content than individual images and the amount of raw data in a video sequence can be huge. Besides, in general, we cannot suppose a prior structure: objects may appear and disappear during the sequence, there is no control of lighting conditions, the object scales cannot be predict in advance, etc. Moreover, video sequences are very redundant and its processing for data retrieval requires a structural analysis to detect shots, extract key frames, and segment scenes. In the Figure 3, these ideas are implemented through a video segment representation (VSR) and optimal key frame representation (OFR). The former is the result of the scene segmentation and simplifies the shot definition. The optimal key frame that represents the video sequence is computed by considering the probability of occurrence of the pixel intensity values with respect to the pixel location among every frame presented in a raw video.

After these simplifications we can perform feature extraction from the segmented video units (shots). The aforementioned steps are applied for the videos in the database as well as for the query video. Finally, a similarity measure is used in order to retrieve the target videos. The scheme presented in the Figure 3 is presented and implemented in [7] using: Expectation Maximization (EM) algorithm and bi-orthogonal wavelet transform for background estimation and video segmentation, respectively; histogram analysis of each pixel along the video sequence to define the optimal key frame; DCT and correlation measure to identify the frames involved in each shot; a general version of the local binary pattern (LBP) operator for texture description and color moments in the Y, Cb, Cr color space; feature weighted distance for retrieval of relevant video clips.

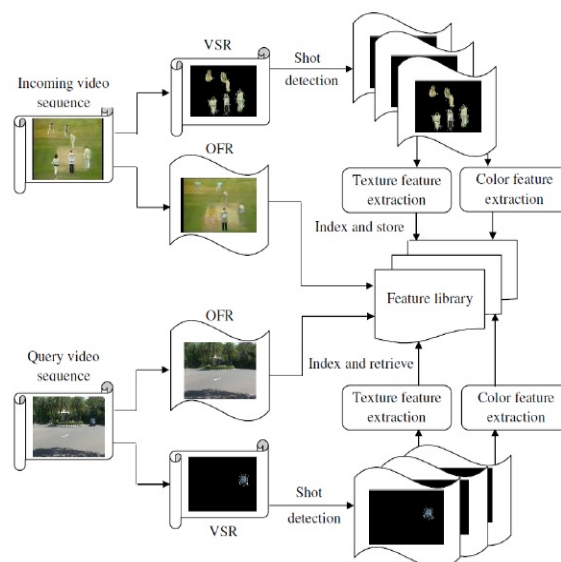


Figure 3: Content-Based Video Retrieval methodology (Reprinted from: [7]).



Up to now, we are in charge with scene segmentation using an alternative methodology. In this strategy, the quadtree is used to select the seed points which are the input for the region growing algorithm. Basically, each leaf of the quadtree is visited and the center of the corresponding region in the image is taken as a seed for the region growing algorithm. After this process it is possible that some

regions of the image remains unsegmented due to inhomogeneity in the image field. A simple search is used to detect these regions which are considered new clusters of the image. We also correct the under and over-segmentation in a post-processing step in order to obtain a fine partition of each frame of the video sequence. Then, we perform a bottom-up strategy based on region merging criteria to define coarse partitions more suitable for object detection. At the end of this process we have completed the video segmentation and object extraction.

The Figure 4 shows a result obtained with this methodology. The Figure 4(a) shows a frame of the video sequence of the [11] In the Figure 4(b) it is pictured the quadtree structure generated with the following parameters: threshold = 0,7 and minimal tableau = 16. The region growing result is shown in Figure 4 (c). We observe over-segmentation of some regions and, on the other hand, there are clusters that include pixels of more than one object (under-segmentation).

We must correct the under and over-segmentation in a post-processing step in order to obtain a fine partition of each frame of the video sequence. A region splitting mechanism should be introduced to separate clusters and correct under-segmentation. This will complete the fine partition of the original image.

The task of over-segmentation reduction can be addressed by a methodology for merging some of the clusters in a proper way. This is performed using a bottom-up strategy based on region merging criteria to define coarse partitions more suitable for object detection. At the end of this process we have completed the video segmentation.

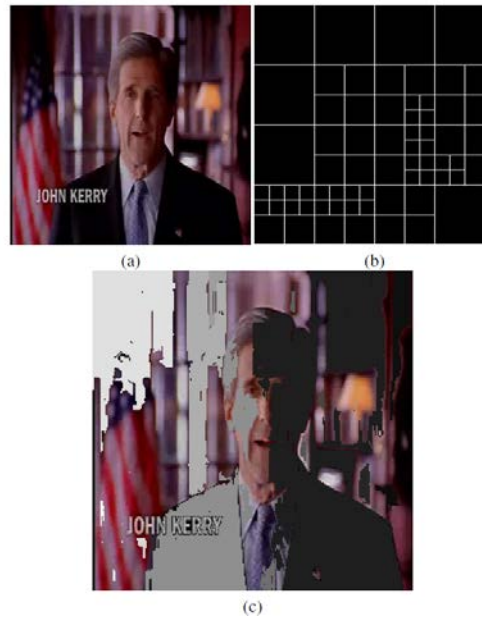


Figure 4 - (a) Original image. (b) Quadtree structure visualization. (c) Segmentation result obtained with the combination of quadtree and region growing.

We will follow the reference [2] to perform both these tasks.

Final Comments

This paper reports our recent work in content-based image and video retrieval. We described the work presented in [1] as well as our proposal to implement the video retrieval pipeline described in [7]. In this way, we expect to obtain a general video retrieval methodology that could be customized for data retrieval in specific databases, like scientific and medical ones.

Acknowledgements

We would like to thank CNPq and CAPES for the financial support for this work.

References

- [1] Douglas Natan Meireles Cardoso, Gilson A. Giraldo, and Luiz Antônio Pereira Neves. Integrando recursos de recuperação de imagens baseada em conteúdo no sistema SBIM. *IEEE America Latina*, Submitted in 2014.



- [2] C.C. Dorea, M. Pardas, and F. Marques. Trajectory tree as an object-oriented hierarchical representation for video. “*Circuits and Systems for Video technology*”, *IEEE Transactions on*, 19(4):547–560, April 2009.
- [3] R.C. Gonzalez. *Digital Image Processing*. Reading Addison-Wesley, 1992.
- [4] Weiming Hu, Nianhua Xie, Li Li, Xianglin Zeng, and S. Maybank. A survey on visual content-based video indexing and retrieval. *Systems, Man, and Cybernetics, Part C: Applications and Reviews*, *IEEE Transactions on*, 41(6):797–819, Nov 2011.
- [5] Anil K. Jain. *Fundamentals of Digital Image Processing*. Prentice-Hall, Inc., 1989.
- [6] L. A. P. NEVES, G. M. PORCIDES, and G. A. GIRALDI. An Internet Manager for Sharing Biological and Medical Images. *Journal of Health Informatics*, 3(4):137–145, 2011.
- [7] S. Padmakala, G. S. AnandhaMala, and M. Shalini. An effective content based video retrieval utilizing texture, color and optimal key frame features. *In Image Information Processing (ICIIP), 2011 International Conference on*, pages 1–6, Nov 2011.
- [8] H. Samet. The quadtree and related hierarquical data structures. *ACM Computing Surveys*, 2:187–259, 1984.
- [9] Vladimir N. Vapnik. *Statistical Learning Theory*. John Wiley & Sons, INC., 1998.
- [10] Ela Yildizer, Ali Metin Balci, Mohammad Hassan, and Reda Alhajj. Efficient content-based image retrieval using multiple support vector machines ensemble. *Expert Syst. Appl.*, 39:2385–2396, February 2012.
- [11] “Trec Video Retrieval Evaluation: TRECVID”, <http://trecvid.nist.gov/>, accessed 15.09.2014.



Estudo da Implementação de Computação Quântica

Aluno: Edgard Winter da Costa

Orientador: Renato Portugal

Introdução

A computação quântica surgiu nos anos 80 com os trabalhos de Benioff e Feynman. Por ser uma área relativamente nova, esperam-se muitas inovações e implementações principalmente nos ramos da física, computação e matemática. A principal linguagem da computação quântica é a álgebra linear, por isso esse estudo teve como principal base análises da álgebra em notação de mecânica quântica.

Objetivo

Os objetivos específicos deste trabalho visam o desenvolvimento da análise de implementação física visando o alcance do computador quântico.

Trabalho Realizado

Até o presente momento já foram estudados todos os princípios básicos da álgebra linear (matrizes de Pauli, decomposição espectral, matrizes Hermitianas, produtos de Kronecker, entre outros), totalmente necessários para o entendimento dos postulados e princípios na mecânica quântica. Foram estudados também os postulados fundamentais da mecânica quântica, medições projetivas e operações com qubits.

Trabalhos Futuros

Como projeto futuro, estudaremos a implementação da física quântica na área de computação quântica com átomos neutros.



Uso de estimativas de precipitação por radar a bordo de satélite na Região Serrana do Rio de Janeiro

Bolsista: Fabio Magest Filgueiras

Orientador: Dr. Pedro Leite da Silva Dias

Coorientador: Dr. Fabiano Gomes

1. INTRODUÇÃO

Atualmente alunos do Projeto Pesquisa de Iniciação Científica do LNCC vem fazendo uma pesquisa em relação ao Uso de estimativas de precipitação por radar a bordo de satélite com ênfase na região serrana. O objetivo principal dessa pesquisa é explorar a variabilidade temporal e espacial da precipitação da região serrana do Rio de Janeiro com as estimativas baseadas no satélite TRMM que tem um radar meteorológico à bordo. Os resultados desse trabalho serão comparados com outra estimativa de precipitação (canal infravermelho do satélite geostacionário GOES), desenvolvido por um outro aluno de Iniciação Científica. A nossa intenção é permitir a conexão, deste trabalho já existente, com o projeto de um modelo de software computacional chamado PAROFIL (Programa de Automatização do Registro de Ocorrências e Filtro de Informações e Logística) aproveitando o estudo e o ampliando ao mesmo tempo.

2. Projeto de Estimativas de Precipitação

O conceito principal do algoritmo de estimativa de precipitação pelo satélite TRMM/3B42 é produzir estimativas de precipitação sobre os trópicos a partir de uma sinergia entre as estimativas de precipitação feitas no canal do infravermelho ajustadas com os modelos de estimativa de precipitação em microondas do satélite TRMM, com uma redução do erro quadrático médio (RMS). Essas estimativas têm uma resolução temporal de 3 horas e resolução espacial de 0,25o x 0,25o englobando as latitudes de 50oS até 50oN. De uma forma simplificada as estimativas do 3B42 podem ser descritas em quatro estágios:

a) As estimativas de precipitação por microondas são calibradas e combinadas;



- b) Estimativas de precipitação no infravermelho são criadas usando a precipitação calibrada no canal das microondas;
- c) As estimativas no IR e no canal das microondas são combinadas;
- d) Uma escala mensal é aplicada.

Dessa forma, cada campo de precipitação é mais ajustado conforme a eficácia da taxa de precipitação no tempo nominal de observação.

3. Projeto Parofil

O projeto PAROFIL surgiu após relatos acerca de obstáculos no tramite de informações entre o atendimento e o envio da equipe de socorro, relatos esses feitos pelo Comandante do 15º Grupamento de Bombeiros Militares de Petrópolis (15º GBM) do Estado do Rio de Janeiro, e também por integrantes da Defesa Civil da cidade de Petrópolis. A partir de uma breve análise foram identificados diversas outras situações de conflito.

Após uma análise dos relatos das instituições citadas, foi constatada a necessidade da criação de um grande modelo de sistema computacional que permitisse a integração de vários serviços de reação (Reativos) ao atendimento emergencial como registro de informações relacionadas as ocorrências atendidas, dados dos veículos das instituições, repositórios de materiais didáticos sobre procedimentos usados nas instituições, possibilitando assim um acúmulo de dados coesos sobre as situações atendidas. Em relação a prevenção de eventos (Pró-ativo) o modelo de sistema deve providenciar mecanismos que permitam o rastreamento das viaturas para maior controle ao envio das mesmas ao destino das ocorrências, ferramentas e filtros para a manipulação dos dados possibilitando a criação de informações em forma de relatórios ou laudos para a análise em decisões de atitudes de prevenção e uma ferramenta que possibilite a comunicação com dispositivos, satélites, pluviômetros, sites, ou seja com qualquer aparelho que possuir capacidade de identificar mudanças climáticas. Esta ferramenta deve ser capaz de ler, analisar e calcular o nível de importância dessas



mudanças e gerar algum tipo de ação preventiva no sistema para que os operadores do mesmo tomem a devida providência.

4.Ferramentas e Metodologia

Para o desenvolvimento deste projeto como o outro envolvido foram estudadas analisadas e utilizadas algumas ferramentas e tecnologias de informática.

No processamento das informações dos satélites foi utilizada a ferramenta GrADS (Grid Analysis and Display System). O GrADS é uma ferramenta computacional interativa para a análise e exibição de dados de ciência da Terra, ele executa um modelo de dados de 4-Dimensões, onde as dimensões são geralmente latitude, longitude, nível e tempo. Cada conjunto de dados está localizado dentro deste espaço de 4-Dimensões através do uso de um arquivo descritor de dados.

Para o cálculo das informações foi usado o Octave, uma linguagem de programação de alto nível, destinada ao tratamento de problemas para computação numérica. Ele pode efetuar cálculos aritméticos com reais, escalares complexos e matrizes, resolver sistemas de equações algébricas dentre outras.

Na construção do modelo de sistema computacional como o nome indica foram aplicados vários conceitos de modelagem de software, como análise de requisitos, estudo de casos, casos de uso, design patterns (padrões de projeto). Um Design Patterns que será utilizado é o MVC (*Model-view-controller*) ou Três Camadas (Camada de apresentação, Camada de negócios "*BLL*" e Camada de acesso aos dados "*DAL*"), este tipo de arquitetura é muito utilizado para sistemas distribuídos, onde os componentes ficam em servidores distintos.

A tecnologia de programação usada será a linguagem de programação *Java*, tecnologia segura e amplamente utilizada, funciona sendo executada em uma camada separada do sistema operacional, em uma máquina virtual, previamente instalada nos clientes, possui um excelente aceitação no mercado multiplataforma.

Modelagem de banco de dados é outra técnica utilizada na construção deste modelo, a aplicação desta técnica presa por evitar redundância de informações no banco de dados de forma a



evitar conflitos e inconsistência na informação obtida através dos dados armazenados. O sistema pode usar duas tecnologias de banco de dados, o MySQL ou Postgre Sql. O MySQL é mais adequado a pequenas e medias transações, já o Postgre Sql é destinado a aplicações com maior volume de transações e armazenamento de dados.

5. Conclusão

O objetivo é integrar uma tecnologia resultado de um estudo anterior a um modelo de software, para aproveitar o projeto e ao mesmo tempo ampliá-lo, e colocá-lo a disposição de serviços que tornem possível a prevenção da perda de vidas em catástrofes provenientes de distúrbios naturais.

6. Referencias

Arkin, P.A. and B. Meisner (1987). The relationship between large-scale convective rainfall and cold cloud over the Western Hemisphere during 1982-1984. Mon. Wea. Rev., 115, 51-74.

Huffman, G.J., Adler, R.F., Bolvin, D.T., Gu, G., Nelkin, E.J., Bowman, K.P., Hong, Y., Stocker, E.F., Wolff, D.B., 2006: The TRMM Multi-satellite Precipitation Analysis (TMPA): Quasi-Global, Multi-Year, Combined-Sensor Precipitation Estimates at Finer Scales. To appear in Journal of Hydrometeorology, June2006.

Use a Cabeça ! Padrões de Projetos (Design Patterns) - 2ª Ed. Revisada - Freeman, Elisabeth; Freeman, Eric - Alta Books;

Utilizando Uml e Padrões - 2ª Edição - Larman, Craig - Bookman;

Princípios de Análise e Projeto de Sistemas com Uml - Bezerra, Eduardo - Campus;

Sistemas de Banco de Dados - 4ª Ed. - Elmasri, Ramez; Navathe - Pearson Education - Br;

Programação Java para a Web - Melo, Alexandre Altair de; Luckow, Décio Heinzelmann - Novatec.



Implementação do campo de força MMFF94 no programa de predição de estrutura de proteínas GAPF

Frederico J. R. Carlos, Gregório K. Rocha, Fábio L. Custódio, Laurent E. Dardenne

Grupo de Modelagem Molecular de Sistemas Biológicos – LNCC

A predição de estruturas de proteínas (PSP) consiste em prever a conformação tridimensional de cada um dos seus átomos, na qual a proteína assume sua função biológica. É uma área de pesquisa de grande destaque nos dias atuais por seu potencial para aplicações em áreas estratégicas de grande impacto na biotecnologia como, e.g., na criação de novas proteínas (engenharia de proteínas), no desenho racional de fármacos, no refinamento de modelos teóricos, na obtenção de estruturas a partir de dados experimentais incompletos e na obtenção e otimização de novas informações dos bancos de dados genômicos. Os principais desafios dos métodos computacionais para PSP estão voltados para a obtenção de possíveis estruturas que a proteína pode adquirir (desenvolvimento de métodos de busca) e na construção de funções de energia que identifique dentro dessas conformações aquela que seria a estrutura encontrada *in vivo* (acredita-se que esta esteja associada ao mínimo de energia). Muitos métodos utilizam os chamados Campos de Força como função de avaliação da energia das estruturas. Os Campos de Força são um conjunto de equações que buscam modelar as principais forças moleculares. Um programa para a predição *ab initio* de estrutura de proteínas chamado GAPF (*Genetic Algorithm for Protein Folding*) vem sendo desenvolvido no Grupo de Modelagem Molecular de Sistemas Biológicos (GMMSB) do LNCC. Apesar dos recentes avanços o método ainda precisa ser otimizado, principalmente no que se refere à função de energia empregada. O campo de forças Merck Molecular Force Field (MMFF94) apresenta uma formulação diferente do atual campo de forças utilizado no GAPF (GROMOS43b1). O MMFF94 é um campo de forças parametrizado usando cálculos quânticos de alta qualidade, podendo ser empregado em um conjunto diversificado de moléculas orgânicas sendo, inclusive, utilizado com sucesso no programa de atracamento molecular DockThor, desenvolvido (GMMSB/LNCC), e em outras abordagens. Assim, este trabalho tem como objetivo implementar o Campo de Forças MMFF94 no programa GAPF visando uma melhoria no seu poder de predição e avaliar o seu desempenho frente ao campo de forças atual.



Computação Quântica

Bolsita: Hugo Souza

Orientador: Renato Portugal

Resumo

Como enganar um computador quântico? Essa máquina poderá ser capaz de reproduzir o pensamento humano?

Chips conseguem fazer a leitura de códigos binários 0 e 1. A coisa está mudando átomos já estão tomando o lugar dos chips. Um computador clássico tem uma [memória](#) feita de [bits](#). Cada bit guarda um "1" ou um "0" de informação. Um computador quântico mantém um conjunto de [qubits](#). Um qubit pode conter um "1", um "0" ou uma sobreposição destes. Em outras palavras, pode conter tanto um "1" como um "0" ao mesmo tempo. O computador quântico funciona pela manipulação destes qubits.

Introdução

Um computador quântico é um dispositivo que executa cálculos fazendo uso direto de propriedades da mecânica quântica, tais como sobreposição e interferência. Teoricamente, computadores quânticos podem ser implementados e o mais desenvolvido atualmente, o en:D-Wave Two, trabalha com 512 qubits de informação. O principal ganho desses computadores é a possibilidade de resolver em tempo eficiente, alguns problemas que na computação clássica levariam tempo impraticável (exponencial no tamanho da entrada), como por exemplo, a fatoração em primos de números naturais. A redução do tempo de resolução deste problema possibilitaria a quebra da maioria dos sistemas de criptografia usados atualmente. Contudo, o computador quântico ofereceria um novo esquema de canal mais seguro. Um computador quântico pode ser implementado com alguns sistemas com partículas pequenas, desde que obedeçam à natureza descrita pela mecânica quântica. Pode-se construir computadores quânticos com átomos que podem estar excitados e não excitados ao mesmo tempo, ou com fótons que podem estar em dois lugares ao mesmo tempo, ou com prótons e nêutrons, ou ainda com elétrons e pósitrons que podem ter um spin ao mesmo tempo "para cima" e "para baixo" e se movimentam em velocidades próximas à da luz. Com a utilização destes, ao invés de nano-cristais de silício, o computador quântico é



menor que um computador tradicional. Uma molécula microscópica pode conter muitos milhares de prótons e nêutrons, e pode ser usada como computador quântico com muitos milhares de qubits. A grande questão a ser resolvida hoje para a implementação destas máquinas é a capacidade de controlar este sistema, já que as interferências são grandes e o tempo de coerência dos estados das partículas, pequeno. Mas este problema já está próximo ao seu fim.

Métodos e Matérias

Na computação clássica o computador é baseado na arquitetura de Von Neumann que faz uma distinção clara entre elementos de processamento e armazenamento de dados, isto é, possui processador e memória destacados por um barramento de comunicação, sendo seu processamento sequencial. Entretanto os computadores atuais possuem limitações, como por exemplo na área de Inteligência Artificial (IA) onde não existem computadores com potência ou velocidade de processamento suficiente para suportar uma IA avançada. Dessa forma surgiu a necessidade da criação de um computador diferente dos usuais que resolvesse problemas de IA, ou outros como a fatoração de números primos muito grandes, logaritmos discretos e simulação de problemas da Física Quântica.

Resultados

A Lei de Moore afirma que a velocidade de um computador é dobrada a cada 18 meses. Assim sempre houve um crescimento constante na velocidade de processamento dos computadores. Entretanto essa evolução pode atingir um certo limite, um ponto onde não será possível aumentar essa velocidade e então se fez necessário uma revolução significativa na computação para que este obstáculo fosse quebrado. E assim os estudos em Computação Quântica se tornaram muito importantes e a necessidade do desenvolvimento de uma máquina extremamente eficiente se torna maior a cada dia.

Discussão



A pesquisa para o desenvolvimento da computação quântica iniciou-se já na década de 50 quando pensavam em aplicar as leis da física e da mecânica quântica nos computadores. Em 1981 em uma conferência no MIT o físico Richard Feynman apresentou uma proposta para utilização de sistemas quânticos em computadores, que teriam então uma capacidade de processamento superior aos computadores comuns. Já em 1985, David Deutsch, da Universidade de Oxford, descreveu o primeiro computador quântico, uma Máquina de Turing Quântica, ele simularia outro computador quântico.

Depois de Deutsch apenas em 1994 houveram notícias da computação quântica, quando em Nova Jersey, no Bell Labs da AT&T, o professor de matemática aplicada Peter Shor desenvolveu o Algoritmo de Shor, capaz de fatorar grandes números numa velocidade muito superior à dos computadores convencionais. Em 1996, Lov Grover, também da Bell Labs, desenvolveu o Speedup, o primeiro algoritmo para pesquisa de base de dados quânticos. Nesse mesmo ano foi proposto um modelo para a correção do erro quântico. Em 1999 no MIT foram construídos os primeiros protótipos de computadores quânticos utilizando montagem térmica. No ano de 2007 surge o Orion, um processador quântico de 16 qubits que realiza tarefas praticas foi desenvolvido pela empresa canadense D-Wave. Em 2011 a D-Wave lançou o primeiro computador quântico para comercialização, o D-Wave One, que possui um processador de 128 qubits. Porém o D-Wave One ainda não é totalmente independente, precisa ser usado em conjunto com computadores convencionais.

Conclusão

De modo geral e simplificado, a computação quântica nada mais é do que a aplicação das teorias e propriedades da mecânica quântica ao computador. Claro que isto envolve muito mais coisas do que podemos imaginar, mas em resumo, é disto que se trata. Um computador quântico (também conhecido como CQ) é capaz de realizar cálculos utilizando propriedades da mecânica quântica, como a teoria da sobreposição. Em teoria, os computadores quânticos são assustadoramente mais rápidos do que o PCs que temos hoje, mas a própria física quântica possui seus limitadores. Um CQ pode ser construído a partir de fótons, nêutrons, prótons, elétrons e até mesmo pósitrons, aquelas partículas presentes no átomo



que você estudou nas aulas de química do colégio. Sendo assim, já é de se imaginar que os computadores quânticos, pelo menos os processadores, possuirão um tamanho mais do que diminuto, em escala atômica. A quantidade de mudanças que seriam causadas pela criação de um computador quântico é enorme. Não há como estimar com precisão quais seriam todas as alterações, mas algumas delas são previsíveis, e sem grandes esforços. O fato de trabalhar com partículas atômicas torna possível a construção de computadores que sejam tão pequenos a ponto de não serem visíveis a olho nu. Enganam-se aqueles que pensaram que a velocidade de processamento é similar ao tamanho. Os computadores quânticos são centenas de milhares de vezes mais rápidos que todos os PCs e notebooks que existem atualmente juntos.

Tal velocidade se dá justamente pela utilização de partículas atômicas em sua construção, uma vez que elas possuem velocidade próxima à da luz, e o processamento das informações poderia ser feito em total paralelismo, e até mesmo em outra dimensão.



Introdução à Computação Quântica e Geração de Números Aleatórios

Bolsista: Igor de Souza Costa /Faeterj

Orientador: Dr. Renato Portugal /LNCC

Coorientador: Me. Pedro Carlos da Silva Lara /LNCC

Objetivo

Este trabalho tem como objetivo estudar e analisar princípios da Computação Quântica, afim de buscar entendimento em seus conceitos, e busca também, estudar aspectos relacionados à geração de números realmente aleatórios através de processos quânticos. De forma geral este trabalho propiciou ao autor, um maior contato com os estudos em Álgebra Linear e com as teorias da Computação e Mecânica Quânticas, cujas disciplinas não são lecionadas no curso de graduação.

Introdução

A motivação e a ascensão na pesquisa da Computação Quântica vem pelo propósito de estudar técnicas que visam melhorar de forma efetiva a velocidade do processamento de dados, tendo compreensão de que a tecnologia baseada em processadores de silício é largamente usada, apesar de já ser considerada defasada. Segundo a Lei de Moore[1], é previsto que a cada 18 meses o número de transistores em um processador seja aumentado em 100%, sem que seja aumentado seu custo de fabricação, porém, atualmente tal lei quase não se atende, e está para ser revogada pelos fabricantes de microprocessadores. Por via dos problemas encontrados na arquitetura atual da computação clássica, físicos, matemáticos e cientistas da computação buscam desenvolver soluções, e uma alternativa proposta é desenvolver e utilizar processadores e algoritmos quânticos, para a otimização de processamento de dados e cálculos complexos, úteis para áreas que exigem tal poder de um processador, como a física quântica, a astronomia, a matemática, inteligência artificial, etc.

O estudo da geração de números realmente aleatórios se deve a utilidade de suas aplicações em, gerações de chaves criptográficas, dinâmicas de partículas, valores randômicos para jogos, e outras diversas aplicações.



Resumo Geral

Neste projeto foram estudados dentro da temática da computação quântica, seus conceitos básicos, Circuitos Quânticos e o Algoritmo de Grover. Ainda há uma previsão e estudos dos números gerados pela placa quântica IDQ-Quantis até o final da vigência da bolsa.

Conceitos Básicos (O Computador Clássico vs O Computador Quântico):

Um computador clássico, descrito genericamente, nada mais é que um dispositivo, que através de entradas de informação em forma de bit, que é a menor unidade de informação que pode ser armazenada ou transmitida, codificada em zeros e uns, faz a leitura de tal entrada de n bits, as implementa fisicamente através de componentes eletrônicos, processados através de portas lógicas, AND, OR e NOT, e retorna um valor de saída de n bits, correspondentes ao valor de entrada.

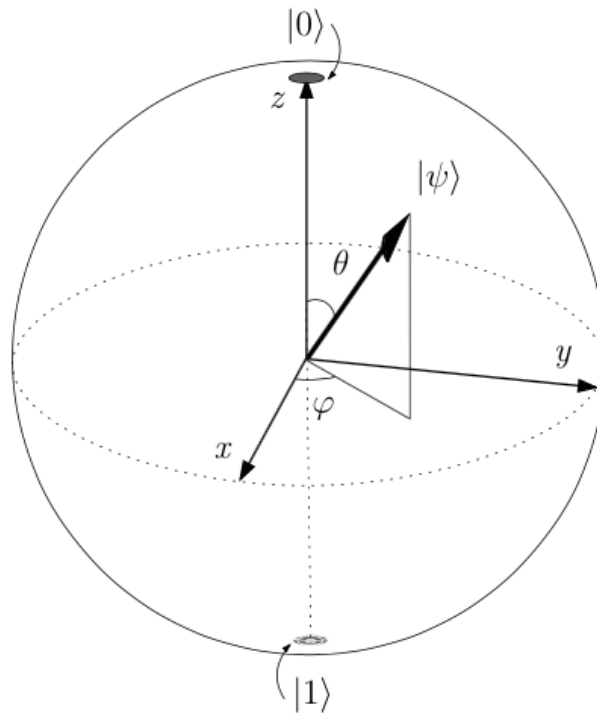
Um computador quântico utiliza estados quânticos de informação, então nesse caso, substituímos o bit pelo q-bit, ou bit quântico, e seus valores de 0 e 1 são substituídos pelos vetores $|0\rangle$ e $|1\rangle$, e suas representações são.

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{e} \quad |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

A diferença entre um bit e um q-bit é que um q-bit genérico $|\psi\rangle$ pode também ser uma combinação linear dos vetores $|0\rangle$ e $|1\rangle$, ou seja,

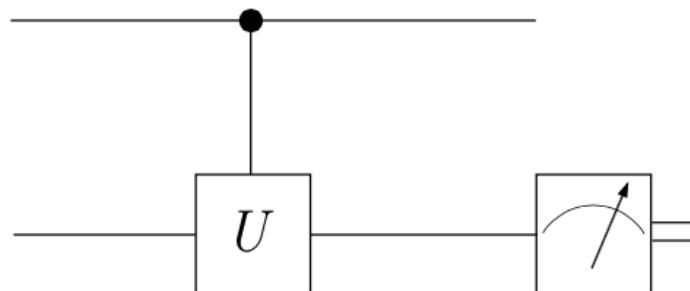
$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle,$$

onde α e β são números complexos. A interpretação física do q-bit acima, é que ele está simultaneamente nos estados $|0\rangle$ e $|1\rangle$. Isso faz com que a quantidade de informação que pode ser armazenada no estado $|\psi\rangle$ seja infinita. Entretanto, essa informação está no nível quântico. Para torná-la acessível, no nível clássico, precisamos fazer uma medida[2].



Representação Geométrica de um qbit na Esfera de Bloch.

A representação gráfica de circuitos clássicos se aproxima bastante da realidade física do circuito fisicamente implementado, no entanto, nos circuitos quânticos, os fenômenos ocorrem de forma diferente.



Exemplo: Porta quântica U-Controlada.

A título de resumo, algumas das portas serão demonstradas.

Porta NOT Quântica

No caso clássico, a porta NOT troca o 1 por 0 e vice-versa. A generalização para o caso quântico é dada por um operador X que satisfaz

$$X|0\rangle = |1\rangle \text{ e } X|1\rangle = |0\rangle.$$

Com isso, verifica-se facilmente que a representação matricial do operador X é dada por[3].

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Porta Hadamard

Uma outra porta de 1 q-bit, largamente utilizada, é a porta Hadamard H , definida pelo operador

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Aplicando H no estado $|0\rangle$, obtemos

$$H|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$

que é uma superposição dos estados $|0\rangle$ e $|1\rangle$, onde a probabilidade de se obter um dos estados, ao se fazer uma medida do estado $H|0\rangle$, é a mesma: 50%. Aplicando o operador H em cada q-bit de um registrador com 2 q-bits no estado $|00\rangle$, temos: [4]

$$\begin{aligned}
 H^{\otimes 2}|00\rangle &= H|0\rangle \otimes H|0\rangle \\
 &= \left(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \right) \otimes \left(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \right) \\
 &= \frac{1}{2} (|00\rangle + |01\rangle + |10\rangle + |11\rangle).
 \end{aligned}$$

Em notação decimal,

$$H^{\otimes 2}|00\rangle = \frac{1}{2} (|0\rangle + |1\rangle + |2\rangle + |3\rangle)$$

generalizando para estados com n q-bits, obtemos:

$$\begin{aligned}
 H^{\otimes n}|0\dots 0\rangle &= (H|0\rangle)^{\otimes n} \\
 &= \left(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \right)^{\otimes n} \\
 &= \frac{1}{\sqrt{2^n}} \sum_{i=0}^{N-1} |i\rangle.
 \end{aligned}$$

O Algoritmo de Grover.

Considere o seguinte problema: temos uma lista não ordenada com N elementos e desejamos encontrar um elemento específico que está na lista. Classicamente, deveríamos testar elemento a elemento. No pior caso possível, precisaríamos realizar N testes. Como veremos, usando as propriedades da mecânica quântica, a quantidade de “testes” necessários para a identificação do elemento procurado será proporcional a \sqrt{N} . Este resultado será obtido usando o algoritmo de Grover.

Para representar matematicamente o problema, vamos supor que a busca será realizada sobre a lista $\{0, 1, \dots, N - 1\}$, onde $N = 2^n$ para algum número natural n, e que a função

$$f : \{0, 1, \dots, N - 1\} \rightarrow \{0, 1\}$$

Definida por:

$$f(i) = \begin{cases} 1, & \text{se } i = i_0, \\ 0, & \text{se } i \neq i_0, \end{cases}$$

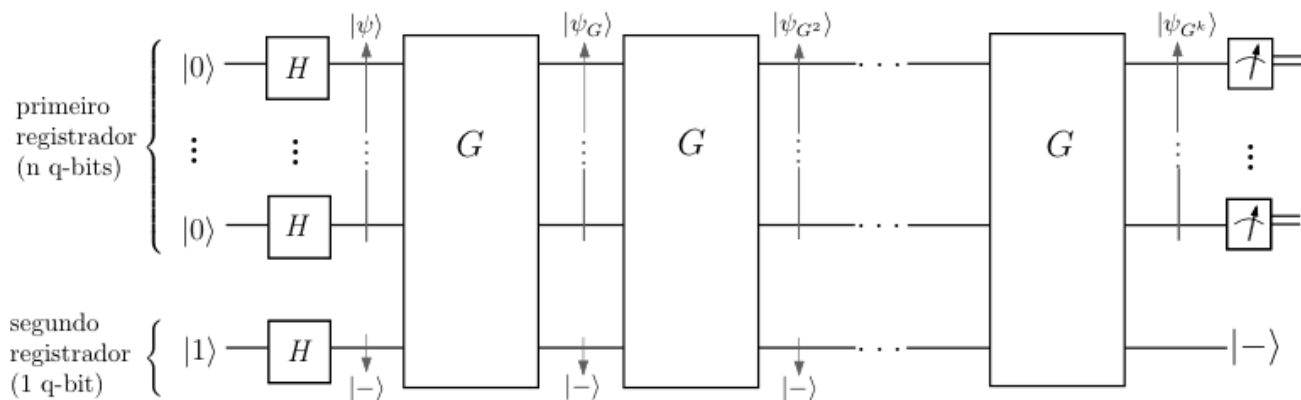
será utilizada para o reconhecimento do elemento procurado i_0 (assumiremos que existe um único elemento $i \in \{0, 1, \dots, N - 1\}$ tal que $f(i) = 1$). Dessa forma, o custo computacional para resolver o problema está associado ao número de vezes que a função f deve ser “utilizada”. Imagine a função f

como sendo um oráculo que está à disposição para informar se um dado elemento é ou não o elemento procurado.

O algoritmo de Grover utiliza dois registradores quânticos: o primeiro, com n q-bits, inicializado no estado $|0\dots 0\rangle$, e o segundo, com 1 q-bit, inicializado no estado $|1\rangle$. O primeiro registrador está relacionado aos elementos da lista onde será feita a busca, enquanto que o segundo é um registrador que terá um papel fundamental, como veremos. A cada elemento i da lista $\{0, 1, \dots, N - 1\}$, associaremos o estado $|i\rangle$ de n q-bits. [5]

Operadores do Algoritmo

Antes da execução propriamente dita do algoritmo, o primeiro registrador é alterado para formar uma superposição de todos os estados associados aos elementos da lista. Isso pode ser obtido aplicando o operador Hadamard H em cada q-bit do primeiro registrador.



Esquema genérico para o Algoritmo de Grover.

Bibliografia

- [1] Gordon Earle Moore, Electronics Magazine 19/04/1965
- [2] Portugal, Renato Uma Introdução à Computação Quântica - São Carlos, SP : SBMAC, 2004 ix, 62 p. - (Notas em Matemática Aplicada; 8) ISBN 85-86883-17-4, pág 14.
- [3] Portugal, Renato Uma Introdução à Computação Quântica - São Carlos, SP : SBMAC, 2004 ix, 62 p. - (Notas em Matemática Aplicada; 8) ISBN 85-86883-17-4, pág 15.
- [4] Portugal, Renato Uma Introdução à Computação Quântica - São Carlos, SP : SBMAC, 2004 ix, 62 p. - (Notas em Matemática Aplicada; 8) ISBN 85-86883-17-4 pág 16.
- [5] Portugal, Renato Uma Introdução à Computação Quântica - São Carlos, SP : SBMAC, 2004 ix, 62 p. - (Notas em Matemática Aplicada; 8) ISBN 85-86883-17-4 pág 22.



CONSTRUÇÃO DE UM SISTEMA DE CONTROLE PARA O CONJUNTO CYBERFORCE

Johnathan Fercher da Rosa, Jauvane C. de Oliveira (Orientador)

Faculdade de Educação Tecnológica do Estado do Rio de Janeiro, Laboratório Nacional de
Computação Científica

1. Dados Gerais

Este documento é uma descrição das pesquisas feitas pelo bolsista Johnathan Fercher da Rosa (CPF 142.382.627-27) para o Programa de Bolsas de Iniciação Científica (PIBIC) do Laboratório Nacional de Computação Científica (LNCC). Os estudos foram supervisionados pelo D.Sc. Jauvane C. de Oliveira.

2. Introdução

A Realidade Virtual [1] é de extrema importância para as mais diversas áreas, pois a recriação de ambientes reais em ambiente simulado possibilita que os usuários tenham contato com alguma situação crítica antes que a mesma venha a acontecer e assim possam estar melhor preparados para enfrentá-las. Um desses Ambientes Virtuais que podemos citar é o *X-Plane* [2] que além de servir como entretenimento traz consigo um modelo de voo realista o bastante para ser utilizado em treinamentos militares [3]. Quando falamos em Ambiente Virtual o nível de imersão é algo que deve ser analisado, pois quanto maior a imersão do usuário, mais caro e mais efetivo será o seu treinamento.

3. Conjunto CyberForce

O *CyberForce* [4] é um conjunto de artefatos criados pela empresa *CyberGlove Systems* no qual o Laboratório A.C.i.M.A [5] vem trabalhando, Dentre os artefatos produzidos pela empresa trabalhamos somente com três, a luva *CyberGlove* que retorna dados da mão humana através de sensores posicionados extrateticamente, o exoesqueleto *CyberGrasp* que fornece uma resistência aos movimentos da mão e o *CyberForce* que rastreia a localização da mão em um ambiente tridimensional (Figura 1).



Figura 1: Na imagem pode-se observar a *Cyber-Glove* vestida, o *CyberGrasp* envolta da mão do usuário e o braço do *CyberForce* conectado aos Artefatos

4. Aplicação

Durante o período de trabalho de Iniciação Científica foi desenvolvido uma classe de controle para a *CyberGlove* e começou-se a desenvolver a interação entre o *CyberGrasp* e a mesma, a aplicação está sendo construída utilizando a linguagem de programação C++ [6] com auxílio das bibliotecas *Glut* [7] e *Bullet* [8]. Na primeira parte do desenvolvimento houve foco na *Cyber-Glove*, pois além de seu funcionamento ser mais simples ela é a base para o funcionamento de todo o resto, para acessarmos os dados da luva foi usado o *SDK* fornecido pela *CyberGlove Systems* e uma classe de acesso aos dados já criada anteriormente.

Para que os dados obtidos da luva sejam confiáveis é necessário que calibremos a mesma, e um dos problemas enfrentados com a calibragem fornecida pelo *SDK* é que além de se levar cerca de trinta minutos para calibrar todos os 22 sensores um a um ela deve ser feita de usuário para usuário, o que a torna extremamente complexa de se lidar (Figura 2).

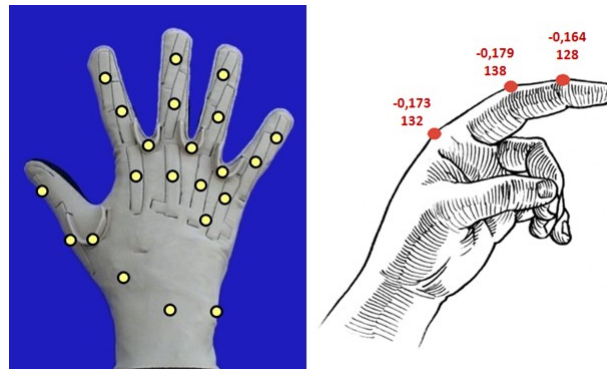


Figura 2: Na imagem a esquerda há a localização de todos os sensores e a direita uma representação dos possíveis valores retornados da luva

Antes que fosse iniciado a implementação de um método de calibragem foram feitos testes para compreender melhor o sistema e, após a constatação de que se tratava de um sistema linear foi criado um método de calibragem visando a simplicidade e robustez onde o usuário deve calibrar os valores dos sensores referentes as posições de mão esticada com o polegar para dentro, completamente aberta, e fechada com o polegar para fora.

Durante a calibragem deve-se tomar cuidado para que a altura das pontas dos dedos nunca ultrapasse a altura do punho (Figura 3).

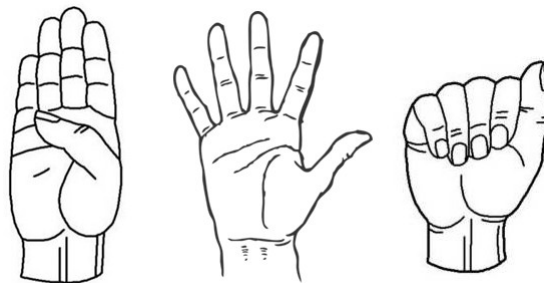


Figura 3: Posições sugeridas para a calibragem da *CyberGlove*



Sabendo a variação dada por um sensor em seus valores máximos e mínimos e a variação angular média de alguma junta da mão dos seres humanos [9] e [10], foi criada uma função onde dado um valor de um determinado sensor temos o retorno do deslocamento angular a partir do ponto zero definido com a mão completamente aberta em radianos, e para que o usuário tenha noções dos resultados foi criado um ambiente virtual simplista com a biblioteca *Glut* que representa a mão humana (Figura 4).

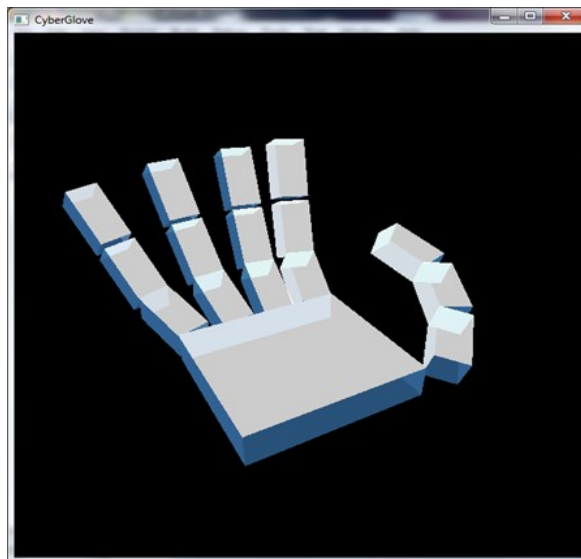


Figura 4: Representação da mão humana em sua posição dado um determinado conjunto de dados retornados da luva

Após obtermos uma representação próxima da movimentação real da mão, começamos os estudos com o *CyberGrasp* que é responsável pelo *feedback* das forças atuantes nas simulações que serão criadas no computador. Para que seja possível fazer com que o exoesqueleto gere uma força oposta ao movimento de abdução da mão e assim passe ao usuário a sensação de estar lidando com algum objeto específico, é necessário antes que saibamos quando a mão virtual colida com o determinado objeto e por isso estamos no processo de implementação da detecção de colisões utilizando a biblioteca *Bullet*.

Para que seja possível recolher os dados da *CyberGlove*, atualizar o modelo virtual e comandar alguma ação com o *CyberGrasp* dentro do menor tempo possível a aplicação foi dividida em uma *Thread* para cada artefato e uma para a representação gráfica dado a diferença de tempo que cada um demora para atualizar e da não necessidade de sincronismo no sistema, assim não nivelando a velocidade do sistema na velocidade de atualização do artefato mais lento.



5. Resultados

O método de calibragem criado para a *Cyber- Glove* é rápido e fácil de ser feito e o sistema fica pronto para utilização em questão de minutos dado o tempo de vestimento da luva, emparelhamento, calibragem, vestimento do *CyberGrasp* e iniciação do programa. O nível de realidade da aplicação virtual é medíocre por tratar-se apenas de um programa de controle não procurou-se ter capricho no modelo 3D e por isso gera pouca imersão no usuário, a calibragem gerada faz com que a representação dos movimentos seja boa, porém é necessário melhorar-la para o polegar e há alguns *bugs* referentes ao cálculo da movimentação em relação aos valores retornados pela luva, o sistema ainda carece da representação de alguns sensores como, por exemplo, os da palma da mão que é representada por um bloco e a adição desses sensores irá adicionar pelo menos duas posições da mão para se calibrar com o método atual.

Foi desenvolvida uma classe de controle para o *CyberGrasp*, porém não está usual, pois a implementação das colisões ainda não está pronta e foi detectado que o jeito que o exoesqueleto é equipado junto a luva, cria alterações significativas nos sensores da mesma.

6. Conclusões

A implementação de um sistema de controle para o conjunto *CyberForce* demonstrou-se ser algo bastante complexo e multidisciplinar, a maior dificuldade de todo o projeto até o momento foi a representação virtual da movimentação dos dedos, que além de utilizar bibliotecas adicionais ao C++ foi necessário um estudo anatômico da mão humana e para o cálculo da representação foram utilizadas algumas técnicas matemáticas até o momento não vistas por mim.

Provavelmente algumas abordagens utilizadas deverão ser mudadas para que alguns problemas sejam resolvidos e para que haja maior precisão no sistema.

Referências

[1] Romero Tori, Claudio Kirner, Robson Siscoutto. "*Fundamentos e Tecnologia de Realidade Virtual e Aumentada*", pp. 9-10 Maio 2006.



- [2] X-Plane. Introduction : What is X-Plane?. Disponível em: <www.x-plane.com>. Acesso em: 12 Setembro 2014
- [3] X-Plane. Professional : FAA Certification. Disponível em: <www.x-plane.com>. Acesso em: 12 Setembro 2014
- [4] CyberGlove Systems. Products: CyberForce. Disponível em: <www.cyberglovesystems.com>. Acesso em: 13 Setembro 2014
- [5] Laboratório A.C.i.M.A. Ambientes Colaborativos e Multimídia Aplicada : LNCC. Disponível em: <www.acima.lncc.br>. Acesso em: 13 Setembro 2014
- [6] Infoescola. Informática : Cpp. Disponível em: <www.infoescola.com>. Acesso em: 14 Setembro 2014
- [7] OpenGL. Libraries : Glut . Disponível em: <www.opengl.org>. Acesso em: 14 Setembro 2014
- [8] Bullet Physics. Disponível em: <www.bulletphysics.org>. Acesso em: 14 Setembro 2014
- [9] PATRÍCIA NETO BARROSO "ANÁLISE DA VARIÇÃO ANGULAR FUNCIONAL DA ARTICULAÇÃO TRAPEZIOMETACARPIANA" Fevereiro 2007
- [10] Danilo A. P. Nagem, Marco A. G. Moreira, Guilherme A. S. Pereira, Carlos J. Tierra-Criollo, Marcos B. Pinotti "Desenvolvimento das Relações Interfalangeanas e metacarpo-falangeanas para os Dedos Durante Movimentos de Pinças" Revista Matéria v. 12, n. 1, pp. 179 – 185, 2007



Reconhecimento de padrões e interpolação de dados com aplicações para imagens médicas

Denise Guliato¹, Gilson Antonio Giraldo², Larissa Alves Pereira¹

1.Universidade Federal de Uberlândia – Faculdade de Computação – UFU

2.Laboratório de Computação Científica - LNCC

Palavras Chave: Interpolação, Segmentação de Imagens, Recuperação de Imagens

Muitos algoritmos têm sido desenvolvidos para reconstrução tridimensional a partir de um conjunto de imagens 2D de seções transversais de um objeto 3D. Muitos destes métodos utilizam informação de contorno da forma a ser reconstruída, criando modelos precisos para regiões de interesse no espaço 3D. Aplicações em medicina envolvendo imagens de tomografia computadorizada e de ressonância magnética frequentemente fazem uso destes algoritmos. Em geral, as seções transversais de um objeto também conhecidas como *slices* não são próximas o bastante, havendo um *gap* de informação entre os *slices*. Assim, uma interpolação se faz necessária, de modo a preencher essa lacuna de informação e reconstruir a imagem desejada.

Algumas abordagens para a interpolação estão sendo analisadas, como em Guo et al. Em [1], que descreve um método de interpolação usando operações morfológicas. Porém o método não consegue interpolar corretamente contornos que possuem formas espiculadas. A abordagem definida por Luo e Hancock [2] possui um resultado considerado suave com relação aos obtidos pelo método de Guo, como demonstra o objeto resultante ilustrado na figura 1. Entretanto, esse método não permite determinar a quantidade de *slices* a serem gerados pela interpolação, o que acarreta em uma imagem distorcida na análise dos resultados uma vez que os objetos *intraslices* não se mantêm em número constante para toda a imagem o que pode vir a ser um efeito indesejável em determinadas aplicações.

São propostas duas novas abordagens para a segmentação e reconstrução tridimensional, a fim de facilitar a aplicabilidade do método para imagens reais em três dimensões. Ambos desenvolvidos segundo pesquisa referenciada em [3], o método SIMOL inspirado na abordagem de Guo, simplifica a implementação, pois utiliza operações morfológicas clássicas e permite a interpolação correta entre formas espiculadas, resultado que pode ser visualizado através da figura 2. Por outro lado, o método

SIDITRANS, inspirado no trabalho de Luo e Hancock, dispensa as operações morfológicas e permite especificar a quantidade de *slices* a serem interpolados.

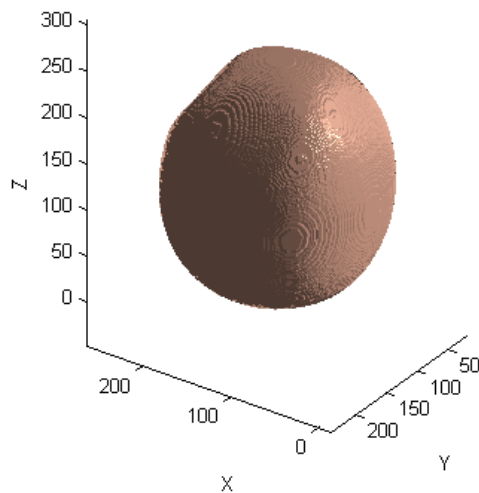


Figura 2 - Objeto em 3D com curvas suaves gerado a partir do algoritmo de interpolação

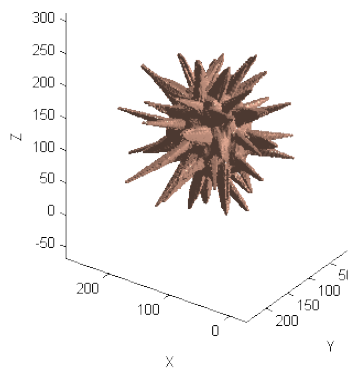


Figura 3 - Objeto em 3D com curvas espiculadas gerado a partir do algoritmo de interpolação



Bibliografia

1. [Guo et al. 1995] Guo, J.-F., Cai, Y.-L., e Wang, Y.-P. (1995). Morphology-Based Interpolation for 3D Medical Image Reconstruction. *Computerized Medical Imaging and Graphics*.
2. [Luo e Hancock 1997] Luo, B. e Hancock, E. (1997). Slice Interpolation Using the Distance Transform and Morphing. *Digital Signal Processing Proceedings*.
3. OLIVEIRA, W. A. A. *HQ feature: descritores de forma utilizando a curva de Hilbert*. 102 f. Tese (Mestrado em Computação), Universidade Federal de Uberlândia, Uberlândia. 2011.



“Estudo de Desempenho de uma aplicação de bioinformática em um ambiente Híbrido de processamento paralelo com tecnologia multicore e GPU”

Mateus Silva de Melo (bolsista), Carla Osthoff (orientadora), Roberto Pinto Souto (coorientador)

Os níveis de proteína nas células são um descritor importante do estado de um organismo. No entanto, como o nível de proteína é muito difícil de ser medido diretamente, é feita uma estimativa indireta utilizando diversas amostras da expressão gênica do organismo, que possui boa correlação com o nível de proteína no organismo. Estas amostras podem apresentar algumas variações nos perfis de distribuição dos valores. Com o objetivo de remover o efeito dessas variações sem origem biológica, é utilizado um método de normalização dos perfis dos valores das amostras, denominado “quantile normalization”, ou simplesmente Q-norm. O maior custo computacional deste método está na etapa de ordenação das amostras. Neste trabalho é mostrado um estudo comparativo entre diferentes implementações de algoritmos de ordenação, executados em CPU (quicksort/insertion sort) e GPU (quicksort e radix sort). Os resultados alcançados, para um estudo de caso com 32 amostras contendo cada uma 6.553.600 valores de expressão gênica, mostraram um efetivo ganho de desempenho computacional com o uso de algoritmos de ordenação em GPU.



Controle e Simulação de Sistemas Sujeitos a Saltos Markovianos

Osacar Neiva Neto; Marcos Todorov

Introdução

Este trabalho é dedicado ao estudo de sistemas sujeitos a saltos Markovianos e questões relacionadas a cadeias de Markov. A pesquisa desenvolvida visa o aprendizado de aspectos teóricos fundamentais relacionados a cadeias de Markov, teoria de controle, e simulação de problemas distribuídos, tais como o cálculo do *PageRank*.

Aplicação

O *PageRank* foi criado pelo Google em 1998 para atribuir valores para páginas de um conjunto de documentos interligados. A ideia por trás do algoritmo é de atribuir pesos às páginas da *world wide web*, em que recebem os maiores pesos as mais visitadas. As principais dificuldades no tratamento deste algoritmo advêm da enorme dimensão do problema e da heterogeneidade da Internet. O algoritmo tem atraído a atenção da comunidade de sistemas e controle, o que se comprova pelo crescente número de artigos relacionados ao tema. Em particular, as referências recentes [2, 5–7] trataram o problema de cálculo do *PageRank* através da teoria de sistemas lineares com saltos Markovianos.

Modelo Matemático

Considere $\theta = \{\theta(k), k = 0, 1, \dots\}$ uma cadeia de Markov homogênea no espaço de estados discreto

$\mathbf{N} = \{1, 2, \dots, M\}$, isto é, tal que:

$$P(\theta(k+1) = j | \theta(k) = i, \theta(k-1), \dots, \theta(0)) = P(\theta(k+1) = j | \theta(k) = i), \quad (1)$$

que é denominada a propriedade de Markov. Tal processo estocástico tem aplicação em diversas áreas da ciência como, por exemplo, computação, engenharia, economia e biologia [1].

O interesse neste trabalho é no estudo de sistemas dinâmicos governados pela seguinte equação:



$$x(k+1) = A_{\theta(k)}x(k) + B_{\theta(k)}u(k), \quad (2)$$

onde x é o estado do sistema, e u representa a ação do controlador.

Os sistemas sujeitos a saltos Markovianos tratam-se portanto de uma classe de sistemas dinâmicos lineares, que possuem parâmetros que se relacionam com o tempo e evoluem de acordo com uma cadeia de Markov em um espaço discreto. O tema é de grande relevância para a teoria de sistemas e controle, uma vez que têm sido uma das principais tendências de pesquisa na área. Existe uma vasta literatura dedicada ao estudo desses sistemas, que sofrem variações abruptas em sua estrutura. Uma amostra significativa de avanços recentes pode ser encontrada nos livros [3,4].

No que concerne a aplicações, este trabalho visa o estudo do algoritmo *PageRank* através de resultados da teoria de sistemas lineares com saltos Markovianos.

Perspectivas Futuras

Neste projeto pretende-se ainda trabalhar com outros métodos de controle e simulação de sistemas estocásticos, que também sejam válidos na simulação do *PageRank*. Aplicações com problemas de consenso de sistemas multi-agentes seriam um desses métodos [8]. Em particular, futuramente pretende-se tratar de métodos baseados em *Markov Chain Monte Carlo* e *Simulated Annealing*.

Bibliografia

- [1] P. Brémaud. *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues*, volume 31 of *Texts in Applied Mathematics*. Springer, New York, 1999.
- [2] K. Bryan and T. Leise. The \$25,000,000,000 eigenvector: The linear algebra behind Google. *SIAM Rev.*, 48(3):569–581, 2006.
- [3] O. L. V. Costa, M. D. Fragoso, and R. P. Marques. *Discrete-Time Markov Jump Linear Systems*. Probability and Its Applications. Springer-Verlag, New York, 2005.



- [4] O. L. V. Costa, M. D. Fragoso, and M. G. Todorov. *Continuous-Time Markov Jump Linear Systems*. Probability and Its Applications. Springer-Verlag, Heidelberg, 2013.
- [5] H. Ishii, R. Tempo, , and E.-W. Bai. PageRank computation via a distributed randomized approach with lossy communication. *Systems Control Lett.*, 61:1221–1228, 2012.
- [6] H. Ishii and R. Tempo. Distributed randomized algorithms for the PageRank computation. *IEEE Trans. Automat. Control*, 55(9):1987–2002, 2010.
- [7] H. Ishii and R. Tempo. The PageRank problem, multiagent consensus, and web aggregation: a systems and control viewpoint. *IEEE Control Syst. Mag.*, 34(3):34–53, 2014.
- [8] W. Ren and R. W. Beard. *Distributed Consensus in Multi-Vehicle Cooperative Control: Theory and Applications*. Communications and Control Engineering. Springer, London, 2008.



Avaliação de desempenho do paralelismo de um problema estreitamente ligado em um ambiente multicore Intel com e sem Hyper-Threading.

Rafael Lourenço Stanislau, Carla Osthoff Ferreira de Barros, Mariano Pereira Silva

O jogo da vida, em termos de paralelismo, é estreitamente ligado assim como os modelos meteorológicos de circulação geral da atmosfera como o OLAM (Ocean-Land-Atmosphere Model). Por isso foi escolhido para teste inicial de desempenho de tecnologias de processamento paralelo visando uso posterior no OLAM. A vantagem do jogo da vida é ser muito mais simples e fácil de alterar. A tecnologia Hyper-Threading da Intel cria na máquina a ilusão de processadores adicionais. Dependendo do problema isso pode ajudar a mascarar latências gerando um ganho de desempenho. O Hyper-Threading na Intel Xeon Phi difere da CPU por ter fisicamente parte do hardware quadruplicado. Este trabalho procurou estudar o desempenho de uma versão paralela do jogo da vida em CPUs Intel com e sem o Hyper-Threading variando a quantidade de threads para futura comparação com a Intel Xeon Phi. Foram utilizados dois computadores, um com processador Intel Core i7 e outro com processador Intel Xeon. Avaliações foram feitas com Hyper-Threading habilitada e desabilitada. No Intel Core i7 o jogo da vida foi executado 5 vezes para situações de 1 a 24 threads, com e sem Hyper-Threading totalizando 240 execuções. Enquanto que, no Intel Xeon foram executadas 10 repetições para situações de 1 a 48 threads com e sem Hyper-Threading totalizando 960 repetições. Menor tempo de execução foi obtido com o número de threads igual ao número de núcleos e sem Hyper-Threading..



Exploração de Ambientes Virtuais com Navegação Natural

Raquel Ellem Marcelino de Oliveira¹ Jauvane Cavalcante de Oliveira²

Palavras Chave: Realidade Virtual, Kinect, Ambiente Virtual, Treinamento, Navegação Natural.

Ambientes Virtuais (AV) com aplicação do mecanismo de *serious games*¹ estão sendo empregados cada vez mais para informar e treinar pessoas sobre situações do mundo real. Treinamentos utilizando AV se tornaram comuns em indústrias de alto valor agregado, como por exemplo, medicina [1], aviação [3] e procedimentos militares [2] diminuindo o custo dos equipamentos necessários e possíveis riscos aos usuários. A segurança conferida por tais ambientes é também um atrativo para esse tipo de treinamento, visto que, se o aluno errar em algum momento, isso não provocará nenhum dano real a si mesmo ou ao equipamento.

A interação é o atributo básico em um AV de treinamento e o usuário deve ser capaz de efetuar os exercícios quantas vezes forem necessárias, de modo que haja aproveitamento das condições proporcionadas pelo ambiente. A possibilidade de treino repetido, o não desperdício de materiais, o reconhecimento do local, a atenuação de custos e a isenção de riscos fazem com que o uso de um sistema deste caráter garanta de forma mais rápida, completa e segura a aprendizagem.

Outro fator relevante é necessidade de acompanhamento contínuo dos usuários, tanto para verificar a evolução no treinamento quanto para corrigir os exercícios feitos de maneira incorreta. Esta situação implica no alto tempo dedicado pelos instrutores e também pelo risco de segurança imposto aos que estão em treinamento, uma vez que podem ocorrer situações para os quais estão despreparados, além da chance de um treinamento incompleto, já que as pessoas não aprendem, necessariamente, no mesmo ritmo. Tendo em vista estas questões, foi desenvolvido um sistema que melhora a qualidade do treinamento, bem como elimina a exposição dos usuários aos riscos do ambiente real. Um sistema deste gênero deverá ser capaz de acompanhar o usuário desde o início do treinamento, até ao momento de conclusão.

O objetivo geral deste trabalho é apresentar um sistema que permita o reconhecimento e treinamento em uma plataforma de petróleo virtual. O recurso utilizado para assegurar a Interação Natural é o Kinect, que rastreia os movimentos do usuário em 3D e tempo real. Com a implementação desenvolvida as informações obtidas pelo dispositivo são recebidas e processadas pelo sistema e o resultado é aplicado no AV, que pode ser visualizado por um monitor ou pelo CAVE, neste último à necessidade de enviar os dados pela rede como mostra a Figura 1.

¹ Este termo passou a ser utilizado para classificar jogos com propósitos específicos, que ofereçam além do entretenimento, experiências voltadas ao aprendizado e treinamento. Eles são destinados a apresentar os problemas do mundo real, visando principalmente situações práticas do dia-a-dia

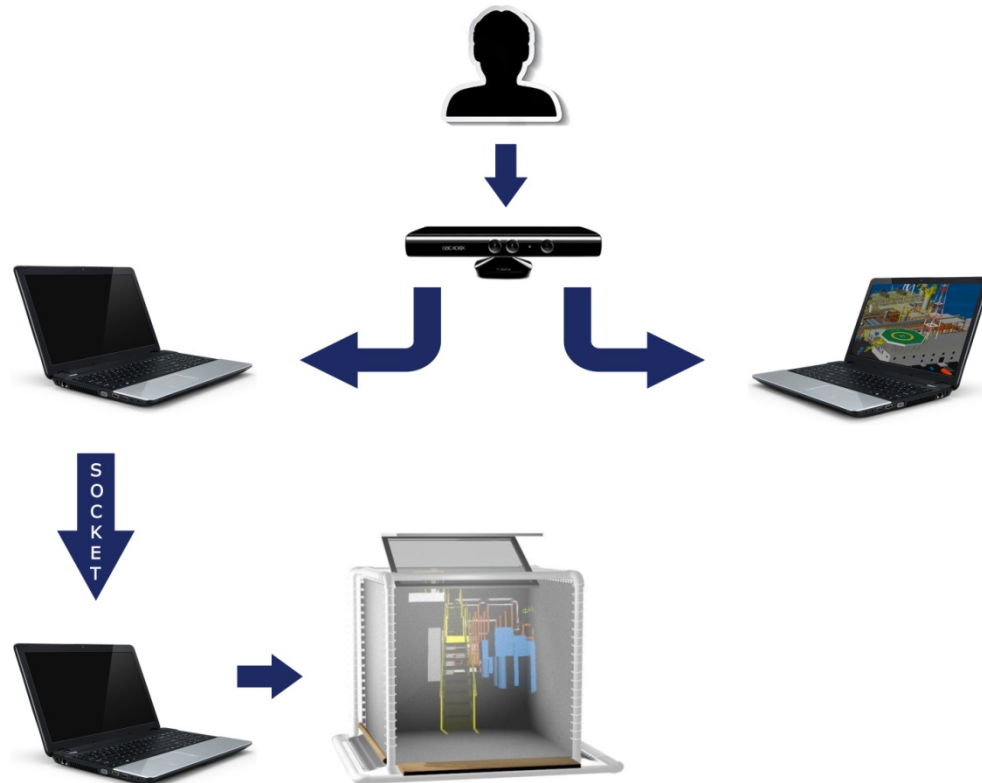


Figura 1 - Modelo geral de funcionamento do sistema. Os movimentos feitos pelo usuário são rastreados em 3D e tempo real pelo Kinect. As informações obtidas são passadas a um *desktop* onde são processadas e o resultado é aplicado no AV. Pode-se visualizar o cenário por um monitor comum ou pelo CAVE

Após receber os valores dos sensores do Kinect, o sistema analisa estes valores baseado em duas possíveis ações do usuário: caminhada ou definição de direção.

A caminhada é definida através da simulação de passos do usuário. Com este movimento a câmera do cenário transladada no eixo Z, fazendo com que o usuário tenha a impressão de estar se movendo para frente.

O usuário também possui a opção de rotacionar a cena. Para isso, deve-se posicionar o braço ao lado do corpo em que se deseja virar. Desta forma, a câmera do cenário será rotacionada para a direita ou esquerda no eixo X. Se o usuário fizer outro ou nenhum movimento, o sistema identifica como parada e por isso a cena visualizada continua a mesma. Por meio destas ações o



usuário pode navegar por toda a plataforma, mudando o ângulo de visão no momento que desejar.

É importante ressaltar que previamente, o usuário é informado sobre a posição inicial e os movimentos básicos a serem realizados, que como visto, tratam-se de ações corporais intuitivas e simples, o que permite às pessoas que nunca tiveram contato com o Kinect executem os movimentos sem dificuldades. Optou-se por manipular a câmera no cenário, a fim de proporcionar ao usuário a sensação de visualizar a cena em primeira pessoa.

As mudanças de cena aqui relatadas são feitas de modo a imitar o olhar humano, sendo assim não há mudanças drásticas no cenário e sim progressivas.

Bibliografia

- [1] F. J. Carter, M. P. Schijven, R. Aggarwal, T. Grantcharov, N. K. Francis, G. B. Hanna, e J. J. Jakimowicz. **Consensus Guidelines for Validation of Virtual Reality Surgical Simulators**. Surgical Endoscopy and Other Interventional Techniques, 19:1523{1532, 2005.
- [2] J. Pair, B. Allen, M. Dautricourt, A. Treskunov, M. Liewer, K. Graap, e G. Reger. **A Virtual Reality Exposure Therapy Application for Iraq War Post Traumatic Stress Disorder**. Em Virtual Reality Conference, 2006.
- [3] L. W. Taylor e K. W. Ili_. **Fixed-Base Simulator Pilot Rating Surveys for Predicting Lateral-Directional Handling Qualities and Pilot Rating Variability**. National Aeronautics and Space Administration, 1969.



VISUALIZAÇÃO ESTEREOSCÓPICA EM MONITORES DESKTOP

Nome: Richard Daniel Calile Salles

Orientador: Jauvane Cavalcante

Aplicações de Realidade Virtual permitem reproduzir o mundo virtual, criado por computador semelhante ao mundo real. Sendo então necessários dispositivos especiais para que o usuário tenha a sensação de presença no ambiente virtual. As aplicações criadas podem ser visualizadas através de monitores comuns de desktops que ainda é o modo mais barato, até por dispositivos especiais como capacetes de RV (Head Mounted Display - HMD) que apresentam a informação através de pequenas telas posicionadas em frente aos olhos do usuário. Outra maneira de visualizar aplicações de Realidade Virtual é a utilização da CAVE, que formada por paredes, teto e piso onde são projetadas imagens estereoscópicas, permitindo assim a imersão do usuário e a visualização do ambiente 3D.[1] [4]

A aplicação exemplo do sistema de visualização estereoscópica, através do uso do monitor, apresenta um modelo 3D no formato 3ds que permite o usuário ter uma noção de profundidade e espaço aumentando assim sua imersão. Assim o presente trabalho oferece suporte a estereoscopia, para aplicações como: jogos, educação e treinamento médico. [2] Aplicações desse tipo tem sido muito favorável para aumentar a motivação para o aprendizado de conteúdo, assim como na área de reabilitação de pacientes com limitação de movimentos. [3]

O presente trabalho apresenta o desenvolvimento de um sistema de visualização estereoscópicas utilizando a API gráfica OpenGL voltadas para monitores de mesa, sendo ainda possível a importação de modelos 3D. O objetivo da aplicação é a exibição de imagens estereoscópicas em monitores desktops, possibilitando a exportação de modelos 3D no formato “.3ds” e ainda permitindo a interação do usuário por meio do teclado. O sistema ainda oferece suporte a futuros trabalho que pretendam utilizar a estereoscopia como ferramenta de visualização, principalmente aplicações de treinamento.

Para que a aplicação seja compatível com o dispositivo 3D Vision, foi necessário a exibição de imagens estereoscópicas a fim de fornecer imagens de um ambiente tridimensional. Além de gerar estereoscopia,



o sistema é capaz também de carregar formatos de arquivos 3D, que amplia as opções de visualização feitas pelo usuário.

O kit 3D Vision contém um sincronizador e uns óculos de LCD, e através da alternância entre as lentes é que a imagem de cada olho é exibida, dando ao usuário final uma imagem estereoscópica. [5] A utilização deste dispositivo possibilita ao usuário imagens em alta qualidade, sem a perda na cor. Para permitir a interação do usuário com a aplicação foi utilizado o teclado como fonte para entrada de dados. Deste modo, é possível interagir com o modelo correspondendo aos comandos do usuário.

Como resultado foi exibido um crânio humano no formato 3ds sendo carregado na aplicação. Visto através do monitor fica notável a exibição das duas imagens devido ao deslocamento entre elas. Porém com o uso dos óculos apenas uma imagem poderá ser vista através das lentes. A visualização dos detalhes do modelo é importante para o usuário que utiliza a aplicação, portanto as funções para zoom também foram implementadas, dessa maneira é possível uma visualização mais detalhada.

REFERÊNCIAS

- [1] CRUZ-NEIRA, C., SANDIN, D. J., DEFANTI, T. A., KENYON, R. V. e HART, J. C. The CAVE: audio visual experience automatic virtual environment. *Commun. ACM*, 35(6):64–72, 1992.
- [2] NUNES, Fátima de Lourdes dos Santos et al.. Realidade Virtual para a saúde no Brasil: conceitos, desafios e oportunidade.
- [3] Machado, L. S. Conceitos básicos da realidade virtual, Monografia, INPE-5975- PUD/025, Instituto Nacional de Pesquisas Espaciais, São José dos Campos/SP, novembro. Disponível on-line em: <http://www.lsi.usp.br/~liliane/conceitosrv.html>, 1995.
- [4] Pimentel, K. & Teixeira, K. Virtual reality - through the new looking glass. 2.ed. New York, McGraw-Hill, 1995
- [5] http://www.nvidia.com/object/quadro_stereo_technology.html Ultimo acesso: 16/08/2014



Modelagem Geométrica e Otimização de objetos 3D para ambientes virtuais colaborativos

Laboratório Nacional de Computação Científica (LNCC)

Stephane Deserto Vasconcelos¹, Jauvane C. de Oliveira²

Palavras Chave: 3D, Otimização.

A implementação de sistemas de Realidade Virtual Colaborativa somente é possível devido a existência de objetos tridimensionais, estes são representados a partir da combinação de polígonos (geralmente triângulos), juntamente com texturas, luzes, etc., a fim de criar representações de uma infinidade de objetos para compor o mundo virtual da forma mais realista possível.

Para a utilização de modelos 3D de alta resolução, é necessário a aquisição de hardwares poderosos e caros, o que acaba por limitar a quantidade de usuários que poderão usufruir dos mesmos. Visando um maior alcance de usuários, o laboratório ACiMA, desenvolve soluções alternativas que utilizam alto desempenho, porém, com custos relativamente baixos.

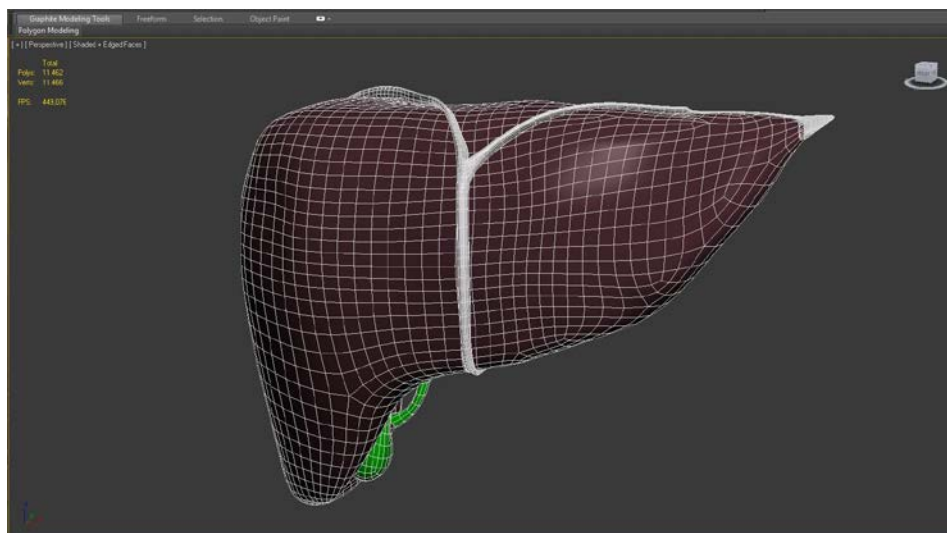
É importante lembrar que só é possível utilizar os modelos nos sistemas atuais de baixo custo, depois que eles passam por um processo de otimização, onde, a otimização da malha, em outras palavras, consiste em diminuir a quantidade de polígonos dos objetos em questão sem que haja grandes alterações em seu visual, tornando assim os modelos geométricos mais "leves" e computacionalmente tratáveis.

Para a criação, manipulação e otimização dos modelos são utilizados softwares de modelagem 3D, tais como, Autodesk 3D Studio Max e Maya.

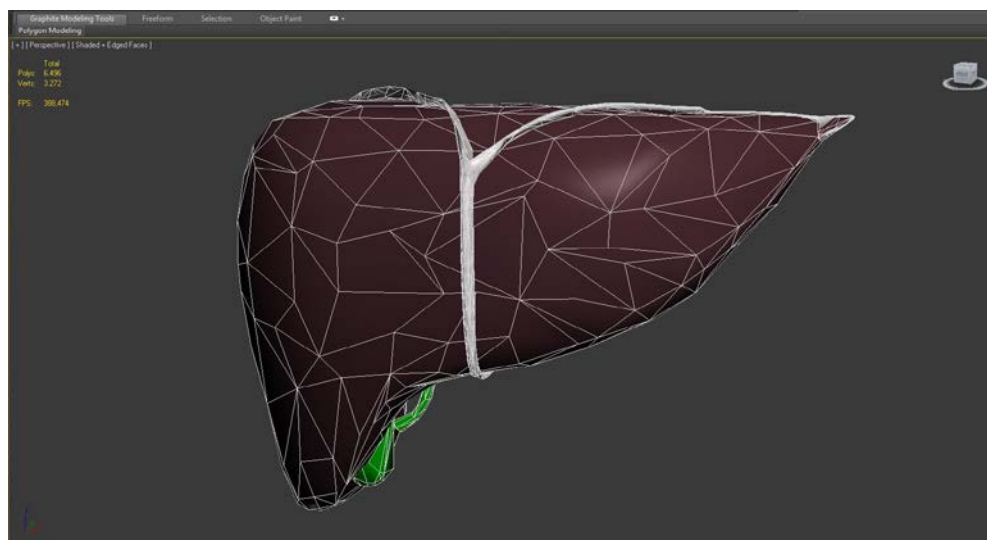
Segue um exemplo comparando uma malha high poly (não otimizada) e uma malha low poly (otimizada), que foi utilizada em um dos projetos do laboratório ACiMA:

Projeto: Atlas anatômico do corpo humano 3D

1.Comparação quanto a quantidade de polígonos:

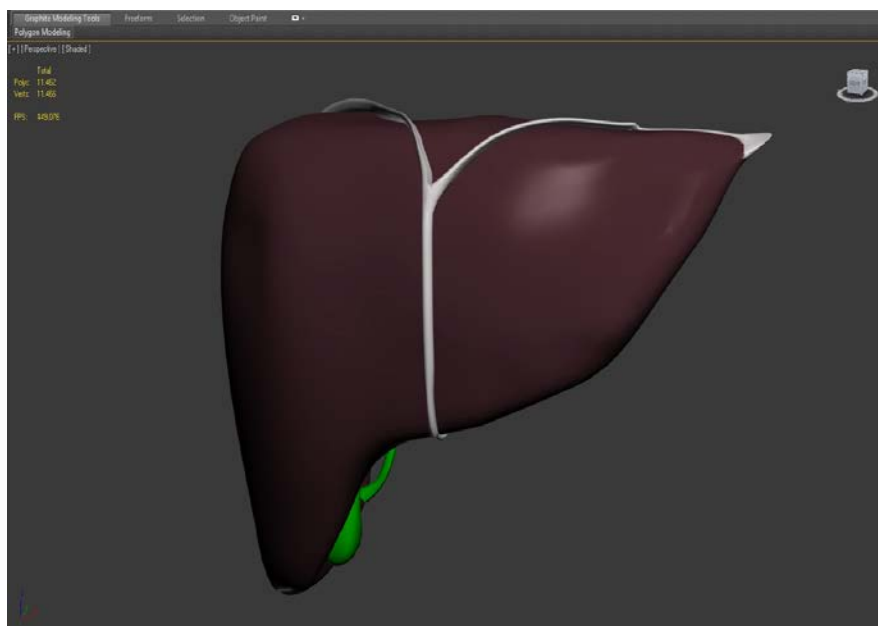


Fígado não otimizado com 11.462 polígonos.

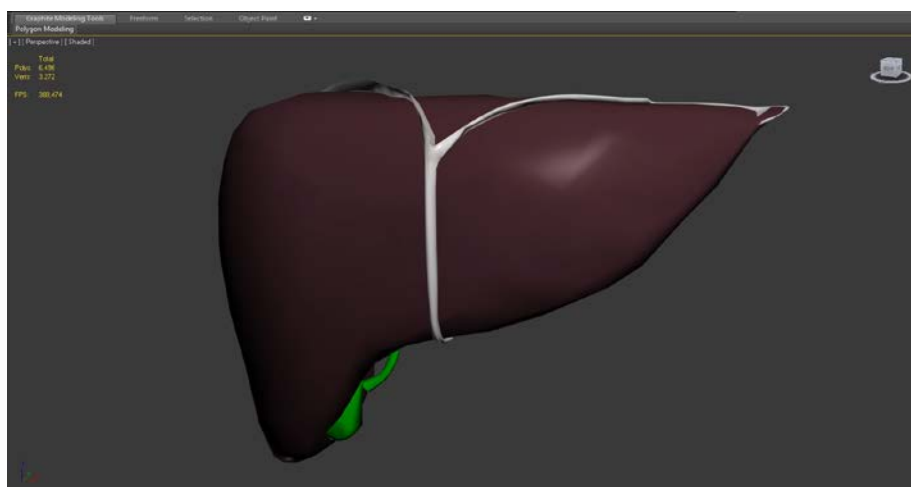


Fígado otimizado com 6.496 polígonos.

Comparação quanto a modificação final da imagem:



Fígado não otimizado com 11.462 polígonos.



Fígado otimizado com 6.496 polígonos.



Modelagem da Manipulação do Comportamento do Hospedeiro –

Caso *Toxoplasma gondii*

Aluna: Tainara Mendes de Andrade Soares

Orientador: Maurício Vieira Kritz

Objetivo

Apresentar um modelo matemático-computacional que represente a manipulação do comportamento do hospedeiro realizado pelo parasita *Toxoplasma gondii*, agente causador da toxoplasmose.

Resumo

A manipulação do comportamento do hospedeiro (MCH) pode ser definida como a influência causada por um parasita ou parasitoide num hospedeiro. Esta influência pode ocorrer em diferentes níveis de intensidade e sempre se expressa como uma modificação do comportamento de seu hospedeiro. Ao contaminar o hospedeiro, o parasito pode induzir o hospedeiro a perfazer ações que o beneficiem em detrimento próprio. Observa-se na maioria dos casos uma influência negativa do parasito sobre o hospedeiro, porém em alguns casos a MCH pode não conferir nenhum valor adaptativo ao hospedeiro.

Hospedeiros parasitados podem apresentar alterações morfológicas, reprodutivas e comportamentais; em função de modificações bioquímicas, fisiológicas e ecológicas (Vinson, 1975; Vinson & Iwantsch, 1980; *capud in* Couto, 2008). A relação parasita-hospedeiro estabelece um conceito no qual as modificações no hospedeiro representam uma extensão do fenótipo do parasito (Dawkins, 1982). Diversos tipos de alterações no comportamento podem ser induzidos, sendo peculiares a cada binômio parasita-hospedeiro. Os comportamentos mais comuns manipulados pelos parasitos são as atividades dos hospedeiros e sua distribuição espacial (Horton & Moore, 1993). A influência do parasito



no comportamento do hospedeiro acontece por vários mecanismos atuantes sobre o sistema nervoso do hospedeiro, através de interações diretas com os sistemas nervoso e muscular, ou via secreção de substâncias neuroativas (Thomas *et al.* 2005). A manipulação comportamental em si sempre ocorre após a liberação de hormônios, neurotransmissores, a injeção de outras substâncias químicas pelo parasito ou o envio de pulsos nervosos por parasitos alojados em locais estratégicos do cérebro ou sistema nervoso central dos hospedeiros.

Neste trabalho estudamos o caso do *Toxoplasma gondii* com o objetivo de representar a modificação do comportamento do hospedeiro intermediário na escala trófica. Este parasita causador da toxoplasmose influencia o comportamento dos ratos, seu hospedeiro intermediário, facilitando a predação destes pelos gatos, seu hospedeiro final. Isto é representado no modelo populacional através da dependência do tempo de captura de ratos infectados por gatos na densidade populacional dos parasitos. Isto acelera o ciclo vital do *Toxoplasma gondii*, visto sua reprodução ocorrer apenas no intestino dos felinos, sendo em seguida dispersos no ambiente. O modelo proposto é contínuo e descreve a dinâmica de quatro populações – parasitas, ratos assintomáticos, ratos infectados e gatos - desconsiderando a dinâmica dos recursos, buscando enfatizar a manipulação do comportamento do rato pelo parasita *Toxoplasma gondii* e sua influência na dinâmica das populações presentes em seu ciclo.



MHOLline 2.0 – A evolução de um *workflow* científico para problemas em Bioinformática e Biologia Estrutural

Victor Crisóstomo Cruz Reis (v.crisostomo.reis@gmail.com)¹, Priscila V. Z. Capriles Goliatt (capriles@lncc.br)^{1,2}, Laurent E. Dardenne (dardenne@lncc.br)²

1 – Departamento de Ciência da Computação – Universidade Federal de Juiz de Fora – UFJF/MG

2 – Grupo de Modelagem Molecular de Sistemas Biológicos – Laboratório Nacional de Computação Científica
GMMSB/LNCC

Resumo

O MHOLline é um *workflow* científico que busca auxiliar em pesquisas nas áreas de Bioinformática e de Biologia Computacional. É composto por um conjunto de *softwares* voltados para a análise sequencial e estrutural de proteínas. As bases para a primeira versão *web* do *workflow* foram apresentadas na tese de doutorado de Shaila C. S. Rössle (2004)^[1]. Em 2010, a parceria com o Laboratório Nacional de Computação Científica proporcionou melhorias ao MHOLline^[2], sendo lançada sua primeira versão *web* no sítio eletrônico www.mholline.lncc.br^[3-4].

O *workflow* foi definido como um ambiente computacional, chamado de MHOL3D, que possui uma estrutura de funcionamento dividida em: (i) MHOLine (código) - responsável pelo processamento dos dados e geração de resultados. Um servidor linux é usado para dar exclusividade no processamento e armazenamento de dados. A linguagem de programação Perl foi escolhida por apresentar ferramentas de processamento de texto avançadas para a época. ShellScript, C e Python são usados para a compatibilização e comunicação entre algumas ferramentas e o servidor propriamente dito; (ii) MHOLweb - plataforma online que interfaceia a submissão dos “processos”, a exibição de seu progresso e os resultados. Apresenta três ambientes de acesso pelo usuário: (a) não logado (com limitações), (b) logado e (c) administrador (possui ferramentas estatísticas, de atualização do sistema e de gerenciamento de usuários). Esta estrutura foi desenvolvida com as linguagens PHP e JavaScript, além de HTML e CSS; (iii) MHOLdb - base de dados usada para armazenamento e consulta de dados de entrada, resultados e controle do sistema. O SGBD usado é o MySQL.

Para a submissão de um novo “processo” o usuário deve inicialmente selecionar os módulos que deseja executar. Algumas ferramentas possuem interdependência, portanto, ao selecionar uma dessas ferramentas automaticamente as outras serão selecionadas. Após a seleção dos módulos, deve-se submeter um arquivo de entrada no formato FASTA, que deverá conter as sequências de aminoácidos a



serem analisadas (e.g., desde uma única sequência até todo um genoma). Caso seja um usuário não logado, este poderá opcionalmente indicar seu endereço de e-mail para que receba um comunicado da conclusão de seu "processo".

O primeiro módulo a ser executado é o BLAST^[5]. Ele busca por proteínas evolutivamente relacionadas, através da comparação do alinhamento local das sequências submetidas com o as que possuam estrutura tridimensional (3D) conhecida e depositada no PDB (Protein Data Bank)^[6]. O programa BATS foi desenvolvido para separar as sequências identificadas pelo BLAST em quatro grupos (G0, G1, G2, G3) de acordo com sua pontuação. Somente as sequências do G2 são destinadas aos próximos módulos do *workflow*. A ferramenta FILTERS foi desenvolvida para classificar as sequências selecionadas pelo BATS em sete grupos distintos, separando de acordo com a qualidade do modelo (de *Very High* até *Very Low*). O programa ECNGet foi desenvolvido para fazer a atribuição de ao menos um número *Enzyme Commission* (EC) para cada sequência a ser modelada cujas proteínas de referência possuam pelo menos uma função enzimática conhecida. A construção do modelo 3D é realizada pelo programa MODELLER^[7], que modela apenas as sequências pertencentes ao G2. Os modelos construídos são avaliados de acordo com sua qualidade estereoquímica pelo programa PROCHECK^[8]. O MHOLline conta ainda com o programa HMMTOP^[9] para a identificação de regiões transmembranares em proteínas e pode ser executado independente dos demais.

Ao final da execução de cada módulo selecionado pelo usuário dentro do *workflow*, serão disponibilizados arquivos com seu resultado padrão (*Output Files*), o resumo de execução (*Resume Files*) e os novos arquivos no formato FASTA (*FASTA Files*). Adicionalmente, o MHOLline gera um arquivo (*Summary*) contendo um resumo geral das informações produzidas em cada etapa do *workflow*.

Neste trabalho, propomos a versão 2.0 do MHOLline. Os objetivos consistem na atualização dos módulos já pertencentes ao *workflow*, adição de novas ferramentas de biologia computacional, realização da padronização e documentação de todo o código, criação de ferramentas para visualização *online* dos resultados e ainda a criação de uma interface de reprocessamento de resultados (selecionados pelo usuário), além da elaboração de uma nova interface *web*.

Até o presente momento, os programas advindos da versão 1.0 já foram atualizados e novos programas já foram acoplados: o MOLPROBITY^[10] (para a validação estereoquímica das estruturas 3D modeladas) e o SIGNALP^[11] (para identificação de regiões de peptídeo sinal em sequências proteicas). Diagramas UML (*Unified Modeling Language*) foram gerados para ajudar no entendimento de cada parte do *workflow*, desde a execução de cada módulo de um novo "processo", ao funcionamento do MHOLweb, passando pelo esquema de tipos de usuários, que possuem permissões e ferramentas diferentes à sua disposição (documentação). A interface *web* do *workflow* foi reescrita e uma nova



interface de visualização de resultados está sendo elaborada, ambas para melhorar a experiência de uso do usuário.

Os próximos passos são: adição de mais ferramentas, com objetivo de expandir o foco da versão 1.0 para a análise e classificação global de proteínas; finalização da interface de reprocessamento e a finalização da padronização e refatoração de todo o código do *workflow*.

Agências de fomento: CAPES e PIBIC/CNPq.

Referências

- [1] Rössle, SCS (2004). **Desenvolvimento de um sistema computacional para modelagem comparativa em genômica estrutural: Análise de seqüências do genoma da *Gluconacetobacter diazotrophicus***. Tese de Doutorado, Instituto de Biofísica Carlos Chagas Filho – Universidade Federal do Rio de Janeiro / Brasil.
- [2] Goliatt, PVZC. (2007). **Técnicas de Bioinformática e Modelagem Computacional Aplicadas ao Estudo do Genoma de *Trypanosoma cruzi* e de Enzimas Consideradas de Interesse no Tratamento da Doença de Chagas: Estudo Particular das Cruzipainas 1 e 2**. Dissertação de Mestrado, Laboratório Nacional de Computação Científica.
- [3] Capriles, PVSZ; Guimarães, ACR; et al. (2010). BMC Genomics, 11:610. doi:10.1186/1471-2164-11-610.
- [4] Guimarães, ACR; Capriles, PVSZ; et al. (2013). In: **iConcept Press Ltd. Genomics II Bacteria, Viruses and Metabolic Pathways**. ISBN: 978-1-480254-145.
- [5] Altschul S.F., Gish W., Miller W., Myers E.W. and Lipman D.J. (1990) **Basic local alignment search tool**. J. Mol. Biol. 215: 403-410.
- [6] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne (2000) **The Protein Data Bank Nucleic Acids Research**, 28: 235-242.
- [7] N. Eswar, M. A. Marti-Renom, B. Webb, M. S. Madhusudhan, D. Eramian, M. Shen, U. Pieper, A. Sali. **Comparative Protein Structure Modeling With MODELLER**. Current Protocols in Bioinformatics, John Wiley & Sons, Inc., Supplement 15, 5.6.1-5.6.30, 2006.
- [8] Laskowski R A, MacArthur M W, Moss D S, Thornton J M (1993). **PROCHECK - a program to check the stereochemical quality of protein structures**. J. App. Cryst., 26, 283-291.



- [9] Tusnady GE, Simon I. (2001) **The HMMTOP transmembrane topology prediction server.** Bioinformatics 17:849-50.
- [10] Vincent B. Chen, W. Bryan Arendall III, Jeffrey J. Headd, Daniel A. Keedy, Robert M. Immormino, Gary J. Kapral, Laura W. Murray, Jane S. Richardson and David C. Richardson (2010) **MolProbity: all-atom structure validation for macromolecular crystallography.** Acta Crystallographica D66: 12-21.
- [11] **SignalP 4.0: discriminating signal peptides from transmembrane regions** Thomas Nordahl Petersen, Søren Brunak, Gunnar von Heijne & Henrik Nielsen Nature Methods, 8:785-786, 2011



Modelagem Computacional Aplicada a Questões Sócio-Educacionais

Aluno: Wesley Peter de Oliveira

Orientador: Dr. José Karam Filho

Co-orientador: Dr. Fabiano Gomes

Resumo

Este projeto de pesquisa tem por finalidade o estudo e o desenvolvimento de ferramentas computacionais para análise tática e estratégica em educação, a fim de auxiliar o processo de tomada de decisão. Até o momento, está em desenvolvimento um sistema para visualização e análise dos dados em ambiente *WEB*. O principal objetivo desse projeto é aplicar técnicas de modelagem e construção de gráficos estatísticos, tendo como motivação questões sócio-educacionais, podendo então generalizar as técnicas e as metodologias à problemas análogos.

A aplicação consiste em um ambiente *WEB* ou seja, seu acesso se dá através de *web browser* (Firefox, Google Chrome, etc), sobre uma rede tal como a internet ou uma intranet. Para o desenvolvimento, é utilizada a linguagem de programação de alto nível, o Java (JSF) e com a utilização do *framework PrimeFaces*, e o gerenciamento do banco é realizado pela ferramenta *PostgreSQL*. Este sistema já permite visualizações de dados do tipo: histograma, *scatter plot*, *box plot* e gráficos de séries temporais, além disto utiliza do *Google maps* para gerar análises. Também permite o calculo de medidas de posição dos dados como média, mediana, variância, desvio padrão, moda, máximos, mínimos e percentis. Técnicas de análise de dados como correlação e componente principal já foram implementadas.

Os dados disponibilizados pertencem a duas unidades de ensino, uma de nível médio e outra de nível superior. Estes dados consistem em notas dos alunos, formação dos professores e alunos evadidos ao longo do tempo.