

Dada uma imagem de consulta, o sistema CBIR busca imagens similares considerando propriedades visuais como a forma

Em busca da imagem perdida

Analista de sistemas desenvolve métodos que ajudam a gerenciar grandes coleções

LUIZ SUGIMOTO
sugimoto@reitoria.unicamp.br

É incomum a abordagem de uma pesquisa de doutorado destacando inicialmente os méritos do autor, mas Daniel Carlos Guimarães Pedronette acaba de conquistar a 3ª colocação no Concurso Nacional de Teses (CTD) promovido pela Sociedade Brasileira de Computação em 2013 e seu trabalho já rendeu 13 publicações em periódicos e conferências, nacionais e internacionais. A tese também foi considerada a melhor do Instituto de Computação (IC) em 2012 e ficou em 2º lugar na 25ª edição do SIBGRAPI, a principal conferência brasileira na área de computação gráfica, processamento de imagens e visão computacional.

Na verdade, os elogios a Daniel Pedronette, cuja ligação com a Unicamp vem desde o curso técnico em informática no Colégio Técnico de Limeira (Cotil), são todos do professor Ricardo da Silva Torres, que orientou o aluno também no mestrado. “Trata-se de um doutorado bastante produtivo em termos de resultados, ainda mais se considerarmos que foi concluído em três anos; e em tempo parcial, já que o aluno seguiu trabalhando como analista de sistemas da Universidade”, observa o docente e atual diretor do IC.

A contribuição mais importante da tese está no desenvolvimento de cinco novos métodos de reclassificação (*re-ranking*) e agregação de listas (*rank aggregation*) visando aumentar a eficácia de Sistemas de Recuperação de Imagens Baseados em Conteúdo (CBIR, na sigla em inglês). Como resultados adicionais, houve a utilização desses métodos para recuperação multimodal (considerando atributos visuais e textuais) e sua execução de forma eficiente a partir do uso de programação paralela e computação heterogênea.

Trocando em miúdos, Ricardo Torres afirma que esta pesquisa contribui para amenizar um problema associado à criação e gerenciamento de grandes coleções de imagens, enfrentado nos últimos anos. “Todos têm hoje a possibilidade de tirar fotos com seus celulares e facilidade de compartilhar o material em redes sociais. Esse grande volume de imagens e de artefatos multimídia de modo geral, demandam, também, novos mecanismos para facilitar o processo de recuperação das imagens de interesse”.

Portanto, ressalta o professor, a grande motivação do trabalho foi criar sistemas de busca mais eficazes. “São dois os norteadores da pesquisa: diante de uma imagem de consulta, o sistema deve devolver o conjunto de imagens mais relevantes em relação ao especificado; há também a questão da eficiência, com o conjunto sendo obtido o mais rapidamente possível. A estratégia escolhida pelo aluno é inovadora por recorrer a diferentes critérios para explorar a informação de ranqueamento – que chamamos de informação contextual.”

O RANQUEAMENTO

Para melhor esclarecer o conceito de ranqueamento, Daniel Pedronette explica que dado um padrão de consulta (a imagem de consulta), um sistema CBIR busca e recupera imagens similares de uma coleção considerando propriedades visuais como cor, textura e forma. “Entretanto, de forma geral, as abordagens de CBIR analisam apenas pares de imagens para a geração das listas de resultados, ignorando informações codificadas nos relacionamentos entre as imagens. Nosso objetivo foi melhorar a eficácia destes resultados fazendo com que cada imagem similar retornada seja tomada como outra consulta. Essa abordagem que explora a informação contextual ao analisar todas as consultas ao mesmo tempo, aproxima-se do comportamento humano que considera as buscas em um determinado contexto.”

Pedronette acrescenta que além da imagem de consulta, outra abordagem comumente utilizada é de palavras-chaves, buscando um texto relacionado e capturando a imagem que está em determinada página. “Essa abordagem por meio de descrições textuais acerca do conteúdo da imagem vem se mostrando um sério desafio diante do crescimento e diversificação das coleções. Ela apresenta problemas de ambiguidade, como por exemplo, quanto à palavra ‘mouse’, que se refere



Fotos: Antonio Scarpinetti

Daniel Carlos Guimarães Pedronette, autor do estudo: “Os métodos servem para qualquer usuário”

tanto ao animal quanto ao dispositivo de computador, não sendo possível a diferenciação.”

Segundo o autor da tese, sua pesquisa também focou a busca multimodal, que combina as abordagens visual e textual para tornar o resultado mais eficaz. “As imagens são classificadas por ordem decrescente de semelhança, de acordo com determinado descritor de imagem. Cada descritor é dotado de um algoritmo para extrair vetores que codificam as características visuais (cor, textura e forma). A similaridade entre duas imagens é calculada em função da distância desses vetores. Demonstramos na tese que os algoritmos propostos podem ser aplicados a qualquer uma das características básicas”.

RANK AGGREGATION

O professor Ricardo Torres ressalta que outra contribuição interessante da tese é a exploração da informação contextual não apenas para melhorar o desempenho do descritor, mas também os resultados da recuperação de imagens a partir da combinação de descritores. “As técnicas propostas permitem melhorar os rankings sobre cor, textura e forma individualmente, e depois combiná-los para gerar um ranking mais relevante em relação às primeiras posições – é o que denominamos agregação de listas (*rank aggregation*) na área de trabalho.”

De acordo com Daniel Pedronette, outra preocupação na pesquisa, além de aumentar a eficácia de Sistemas de Recuperação de Imagens Baseados em Conteúdo, foi a obtenção de bons resultados em tempo hábil. “Hoje as máquinas tendem a migrar para a paralelização da computação. Parte do meu trabalho envolveu a paralelização dos algoritmos visando ao uso mais eficiente de placas gráficas (GPUs) para que os resultados de busca sejam obtidos o mais rápido possível.”

O autor da tese informa que a avaliação experimental por ele conduzida considerou diferentes descritores e coleções de imagens. “Várias propostas têm sido propostas para CBIR, considerando aplicações em áreas diversas como de recuperação de imagem facial, sistemas de informação de biodiversidade, aplicações médicas e imagens de sensoriamento remoto. Os resultados dos experimentos demonstram a eficácia dos métodos que estamos propondo”.

Ricardo Torres não vê restrições em termos de aplicações, como na implementação efetiva desses métodos no sistema de busca de uma biblioteca digital ou de imagens na web. “Eles servem para qualquer usuário, pois é um processo proposto para ser realizado off-line: usa os algoritmos, potencializa a melhora dos rankings e, havendo uma nova consulta, explora essa informação contextual e fornece resultados relevantes o mais rápido possível. Portanto, serve para qualquer público, dependendo apenas da aplicação implementada no sistema de busca.”

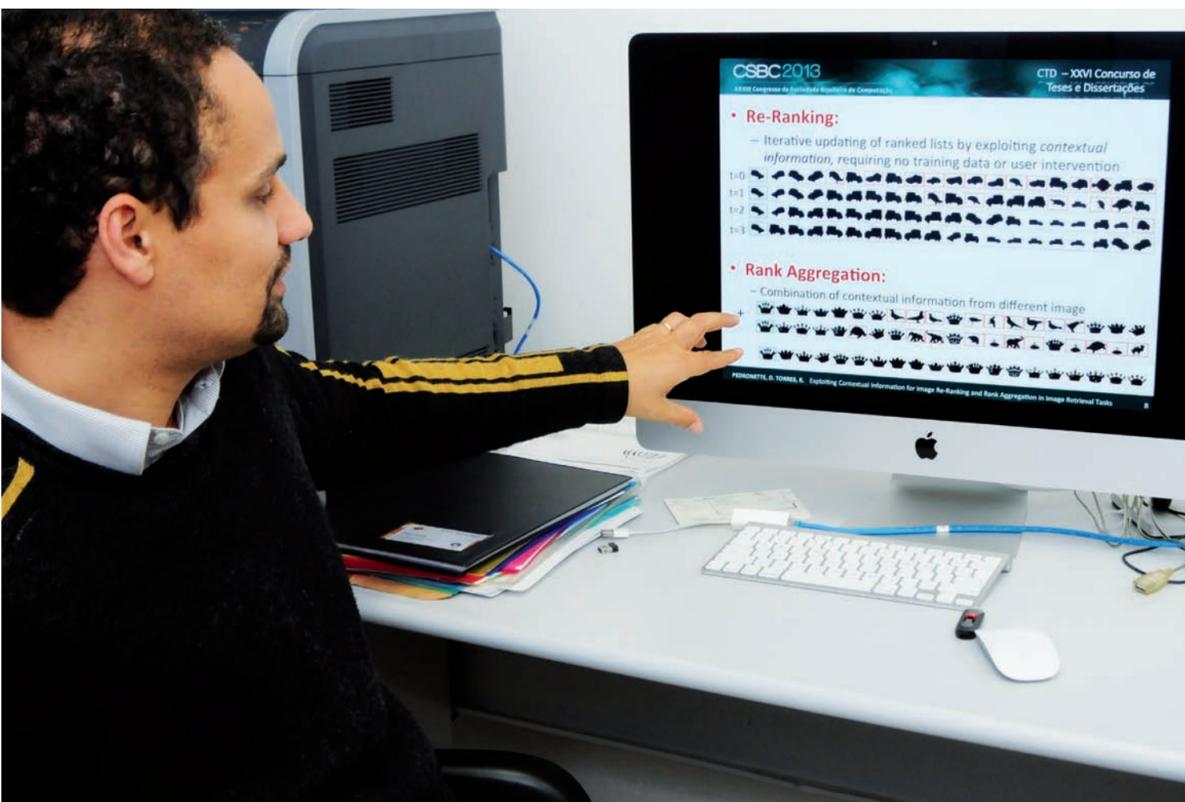
Publicação

Tese: “Exploiting Contextual Information for Image Re-Ranking and Rank Aggregation in Image Retrieval Tasks”

Autor: Daniel Carlos Guimarães Pedronette

Orientador: Ricardo da Silva Torres

Unidade: Instituto de Computação (IC)



O professor Ricardo da Silva Torres, orientador: “Trata-se de um doutorado bastante produtivo em termos de resultados”