

Uma Abordagem Oportunista para a Sugestão de Anotações de Fotos Georeferenciadas

Tales Parente Feitosa
Mestrando

Programa de Pós-Graduação
em Ciência da Computação
Universidade Federal do Ceará
60455-760 Fortaleza, CE, Brasil
tales.parente@sti.ufc.br

Fernando A. Mota Trinta
Orientador

Departamento de Computação
Universidade Federal do Ceará
60455-760 Fortaleza, CE, Brasil
fernandotrinta@great.ufc.br

Windson Viana de Carvalho
Coorientador

Instituto Universidade Virtual
Universidade Federal do Ceará
60455-760 Fortaleza, CE, Brasil
windson@virtual.ufc.br

RESUMO

A massificação de dispositivos móveis e a facilidade de produção de conteúdo multimídia (e.g., vídeos, fotos) proporcionada por esses dispositivos apresentam alguns desafios. Entre eles, o gerenciamento e a recuperação desse tipo de conteúdo. A associação de anotações a esse tipo de arquivos é uma das abordagens que vêm sendo utilizada para o enfrentamento desses desafios. Este trabalho propõe uma infraestrutura de suporte à anotação de fotos capturadas com dispositivos móveis. A infraestrutura proposta dá suporte à anotação de fotos em dois momentos: logo após a captura da foto, quando são sugeridas anotações geradas por algoritmos de processamento mais rápido, e no momento em que o usuário for acessar suas fotos em seu computador pessoal, quando são sugeridas anotações geradas por algoritmos de processamento mais complexo, que executam de forma oportunista e transparente para o usuário.

Palavras-chave

Anotação de fotos georeferenciadas, sensibilidade ao contexto, processamento oportunista.

1. CONTEXTO TEÓRICO

O advento e a popularização de dispositivos móveis com recursos fotográficos são fatores que contribuíram para um fenômeno recente: a produção massiva de conteúdo fotográfico digital por pessoas que querem registrar os mais diversos momentos de suas vidas. As redes sociais, como *Facebook*, *Instagram*, *Flickr* e *Picasa*, também têm contribuído para esse fenômeno, já que fornecem múltiplas formas para o armazenamento e compartilhamento desses registros.

Um dos desafios apresentados por este cenário é o do gerenciamento dessas coleções de fotos pessoais. A associação de metadados às imagens, na forma de anotações, é uma maneira de facilitar esse gerenciamento. A literatura apresenta trabalhos que propõem diversas estratégias de produção de anotações [1][2][3][4][6][8][9][10]. Dentre estes trabalhos, destacam-se aqueles que propõem a utilização de informações provenientes do contexto do usuário quando da captura de uma foto [3][4][6]. Nessas

abordagens, o georeferenciamento da foto é o ponto de partida principal.

Neste trabalho, estamos particularmente interessados no problema da geração automática de anotações que sejam semanticamente relevantes, utilizando dados do contexto. Por anotação semanticamente relevante, entendemos a anotação que descreve bem, do ponto de vista do usuário, algum aspecto (e.g., espacial, temporal, social, conteúdo) da imagem representada em uma foto.

1.1 Anotação de fotos digitais

Para obter os resultados a que se propõem, as ferramentas de gerenciamento de fotos necessitam extrair informações das imagens. Isto é feito, basicamente, a partir de duas fontes: o conteúdo (e.g., *pixels*, cor, textura, formas) ou os metadados associados a elas.

Por derivarem informação de elementos de baixo nível, as abordagens que se baseiam no conteúdo da imagem são mais sensíveis ao problema do distanciamento semântico (*semantic gap*) que existe entre dados brutos, como cor e textura, e conceitos semânticos percebidos por seres humanos, como objetos, pessoas e lugares. Deste modo, estas abordagens apresentam desvantagens como a baixa precisão e a limitação no número de conceitos que podem ser reconhecidos [1].

Outra abordagem é a que utiliza as anotações associadas às fotos. Existem diversas formas de associar anotações às imagens. Por exemplo, alguns formatos de imagem (e.g., JPEG) possuem cabeçalhos específicos (e.g., EXIF, IPTC) que incorporam as anotações ao arquivo da foto, geralmente no formato de pares chave-valor. Em outras estratégias, como nas baseadas em ontologias, as anotações são mantidas em arquivos separados, mas vinculadas às fotos de alguma maneira [6].

1.2 Sensibilidade ao contexto

Os dispositivos móveis atuais são equipados com uma variedade de sensores que permitem a captura de informações do contextuais, que podem ser usadas como anotações preliminares. Dey e Abowd [7] definem contexto como “qualquer informação que pode ser usada para caracterizar a situação de uma entidade e entidade é uma pessoa, lugar ou objeto que é considerado relevante para a interação entre o usuário e

a aplicação, inclusive o usuário e a aplicação”.

Essas anotações preliminares podem servir como ponto de partida para a geração de outras anotações, com maior relevância para o usuário. Por exemplo, a partir da posição geográfica do usuário, que é automaticamente anotada na foto por grande parte dos dispositivos móveis atuais, podem ser derivadas outras anotações, como o nome do lugar referente àquela localidade. Essa capacidade de sistemas de utilizar informações contextuais com o objetivo de melhorar seu comportamento ou de fornecer serviços e conteúdos mais adequados aos usuários é definida como Sensibilidade ao Contexto.

Especificamente na área de gerenciamento de conteúdo multimídia, a literatura apresenta várias soluções sensíveis ao contexto. Por exemplo, Rost *et al.* [10] propõem uma ferramenta de suporte à exploração de fotos baseada em localização. A aplicação captura a localização do usuário à medida que este se desloca e sugere, para visualização, fotos geograficamente relacionadas com a posição do usuário. Os trabalhos propostos em [4] e [6] apresentam abordagens de anotação automática de fotos digitais que utilizam elementos do contextuais como ponto de partida para a geração de anotações semanticamente enriquecidas.

1.3 Geração automática de anotações

Quando o número de fotos aumenta, a anotação manual de fotos apresenta desvantagens. Em [1], Chai *et al.* destacam que a anotação manual é uma tarefa tediosa e que consome bastante tempo. Ma *et al.*[2] apontam que outra desvantagem é a introdução de ruído no processo de anotação manual. O autor explica que ruído, na anotação manual de fotos, é a existência de anotações semanticamente inexpressivas, por serem demasiadamente genéricas.

Dessa forma, diversos trabalhos têm proposto técnicas semi-automáticas ou automáticas de geração de anotações. Dentre as técnicas e algoritmos utilizados, destacam-se: uso de ontologias e regras de inferência, métodos de classificação de conteúdo e algoritmos de recorrência.

No trabalho proposto por Viana *et al.* [3], por exemplo, os autores fazem uso de tecnologias da *web* semântica (ontologias e regras de inferência) para dar suporte à geração automática de anotações. Elementos espaciais (*e.g.*, localização geográfica), temporais (*e.g.*, data e hora) e sociais (*e.g.*, identificação de dispositivos móveis próximos ao usuário) do contexto do usuário são capturados e automaticamente associados às fotos na forma de instâncias das classes da ontologia previamente definida. Num segundo momento, regras SWRL definidas pelo usuário são utilizadas em conjunto com os conceitos da ontologia e com as anotações contextuais para realizar inferências e gerar novas anotações.

Figueirêdo *et al.* [9] propuseram uma ferramenta para gerenciamento de fotos digitais que fornece algoritmos para a geração automática de anotações relacionadas a pessoas e eventos representados nas fotos. Para a sugestão de nomes de pessoas, por exemplo, são

utilizados algoritmos que levam em consideração uma base de anotações prévias e calculam métricas de recorrência (global, espacial, temporal) para se chegar aos nomes de pessoas com maior probabilidade de estarem presentes numa foto. Para exemplificar, seja uma foto F, capturada numa determinada localidade. Dispondo de uma base histórica de anotações, o sistema pode verificar quais são as pessoas que mais vezes foram anotadas em outras fotos capturadas na mesma localidade de F. Essas pessoas terão maior probabilidade de estarem presentes também na foto F. Esse é um exemplo de utilização de métrica de recorrência espacial.

2. IDENTIFICAÇÃO DO PROBLEMA

Com a evolução dos dispositivos móveis, os usuários são capazes de interagir rapidamente com as fotos capturadas, por exemplo, adicionando anotações às fotos. No entanto, esses dispositivos ainda carecem de recursos suficientes para executarem algoritmos mais sofisticados de geração de anotações. Deste modo, uma solução que combine as possibilidades proporcionadas pelos dispositivos atuais (*e.g.*, captura e compartilhamento de fotos, aquisição de contexto por meio de sensores) com o uso de recursos computacionais mais robustos para a execução de algoritmos mais sofisticados de geração automática de anotações se faz necessária.

3. OBJETIVO

O objetivo deste trabalho é propor uma infraestrutura que dê suporte à anotação imediata e *a posteriori* (oportunista) de fotos capturadas por dispositivos móveis. Pretende-se implementar vários algoritmos de geração de anotações na forma de um serviço remoto. Algoritmos mais simples serão utilizados para a sugestão de anotações ainda no dispositivo móvel. Algoritmos mais sofisticados executarão de forma oportunista e serão utilizados para a sugestão de anotações *a posteriori*, por exemplo, quando o usuário descarregar ou acessar as fotos no seu computador pessoal.

4. CONTRIBUIÇÕES ESPERADAS

Espera-se, com a implementação da infraestrutura proposta, fornecer uma solução de suporte à anotação de fotos capturadas por dispositivos móveis com as seguintes características:

- i) sugestão de anotações preliminares, no momento da captura da foto, com anotações geradas por algoritmos mais simples, potencializando, assim, a interação do usuário com o conteúdo fotográfico produzido;
- ii) enriquecimento das anotações preliminares por meio de algoritmos mais sofisticados, fazendo-se uso de processamento oportunista, que acontece de forma transparente para o usuário em infraestrutura remota (*webservice* ou serviço em nuvem). A fase oportunista da sugestão de anotações estaria acoplada a um serviço em nuvem de compartilhamento de arquivos (*e.g.*, *Dropbox*, *Google Drive*);
- iii) utilização de elementos (espaciais, temporais, sociais) do contexto do usuário como ponto de

partida para a geração de anotações semanticamente relevantes para o usuário.

A Figura 1 apresenta a visão geral da arquitetura esperada para o sistema. Essa arquitetura é composta de quatro componentes básicos: a aplicação móvel, a aplicação *desktop*, a infraestrutura de anotação e a infraestrutura de armazenamento. A aplicação móvel executará no dispositivo do usuário e fornecerá funcionalidades para a realização de anotações preliminares nas fotos capturadas. A aplicação *desktop* executará no computador pessoal do usuário e proverá funcionalidades para a validação de anotações adicionais. A infraestrutura de anotação é composta por um serviço remoto (serviço em nuvem ou na forma de *webservices*), que proverá um conjunto de algoritmos de geração de anotações, e uma base de dados, que armazenará informações a respeito das anotações validadas pelo usuário. Essas informações podem ser usadas como entrada para um ou mais algoritmos do serviço remoto para a inferência de novas anotações. A infraestrutura de armazenamento é composta por um serviço de compartilhamento de arquivos na nuvem (*e.g.*, *Dropbox*, *Google Drive*), que faz, de forma transparente para o

para o usuário e já é um serviço comum a várias plataformas de *smartphones* (*e.g.*, *Android*, *iOS*). As setas indicadas pelo número 3 representam a comunicação entre o serviço remoto de geração de anotações e o serviço de armazenamento na nuvem. À medida que novas fotos (que já foram preliminarmente anotadas) forem transferidas para o serviço de armazenamento na nuvem, o serviço remoto inicia a execução de algoritmos mais sofisticados de geração de anotações, utilizando, como entrada, as anotações já existentes nas fotos e outras fontes de dados (*e.g.*, redes sociais, servidores de nomes). Vale ressaltar que a infraestrutura aqui proposta deve se utilizar de serviços de notificação do serviço de armazenamento em nuvem, para iniciar esse processo de forma oportunista. Algoritmos utilizados nessa fase podem usar técnicas que demandam mais tempo e recursos de processamento como, por exemplo, clusterização, reconhecimento de padrões e reconhecimento de faces. A seta indicada pelo número 4 representa a transferência das fotos (já processadas pelos algoritmos mais sofisticados do serviço remoto) do serviço de armazenamento na nuvem para o computador pessoal do usuário. Neste momento, o usuário poderá validar as novas anotações geradas pelo

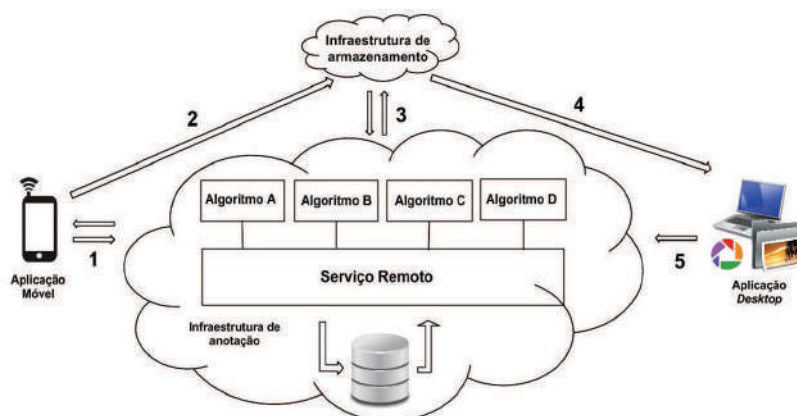


Figura 1: Visão geral da arquitetura

usuário, a transferência e sincronização das fotos entre seus diversos dispositivos (*e.g.*, entre seu *smartphone* e seu computador pessoal).

As setas indicadas pelo número 1 representam a comunicação entre a aplicação móvel e o serviço remoto de geração de anotações. Após a captura de uma foto, a aplicação móvel envia informações do contexto do usuário (*e.g.*, localização geográfica) para o serviço remoto, que utilizará essas informações como entrada para algoritmos de geração de anotações de execução rápida. A partir da localização geográfica, por exemplo, esse algoritmos poderiam sugerir, como anotações, nomes (*e.g.*, cidade, bairro, rua) referentes à respectiva localidade. As anotações geradas são enviadas de volta para o dispositivo móvel para validação por parte do usuário. Uma decisão de projeto que deve ser mencionada é que o usuário tem a opção de ativar ou não esse serviço de anotação imediata. A seta indicada pelo número 2 representa a transferência das fotos capturadas para um serviço de armazenamento de arquivos na nuvem. Essa transferência é feita de forma transparente

serviço remoto. A seta indicada pelo número 5 representa a comunicação entre a aplicação *desktop* de validação de anotações e o serviço remoto. Informações a respeito das anotações validadas pelo usuário são enviadas para o serviço remoto, que as armazena para futura utilização e inferência de novas anotações (*e.g.*, para uso em algoritmos que utilizam métricas de recorrência).

5. METODOLOGIA ADOTADA

A metodologia prevista para a construção da solução proposta consistirá i) da construção de um arcabouço teórico acerca dos temas relacionados à pesquisa (Seção 1), em especial o estado da arte da geração automática de anotações para fotos, ii) do desenvolvimento da arquitetura proposta na Seção 4 e iii) da avaliação da solução proposta.

Para a avaliação, será utilizada uma metodologia baseada em *Gold Standard*. Para a construção do *Gold Standard*, uma base de anotações deverá ser gerada com o auxílio de usuários reais, que serão solicitados a anotar as fotos de uma determinada coleção de forma manual. Em

seguida, a infraestrutura prosta neste trabalho será utilizada para gerar sugestões de anotações para a mesma coleção de fotos e a qualidade das sugestões será avaliada, confrontando-as com o *Gold Standard*.

6. ESTADO ATUAL DO TRABALHO

No estágio atual do trabalho, a revisão de literatura encontra-se finalizada. No entanto, uma revisão de literatura continuada deve ser executada em paralelo às demais atividades previstas. Os padrões de anotação de fotos EXIF e XMP-IPTC foram os escolhidos por serem os mais difundidos na indústria e por já existirem APIs que os implementam. Os primeiros algoritmos do serviço remoto foram identificados, que são: a utilização de fontes externas para agregação de dados [5], os algoritmos que usam métricas de recorrência (global, espacial, temporal) para sugestão de nomes de pessoas [9], e a utilização de ontologias e regras de inferência para enriquecimento de anotações contextuais [8]. Uma aplicação móvel para anotação de fotos (plataforma *Android*) já se encontra implementada [11]. Foram feitas provas de conceito para testes de integração com serviços externos, como *GeoNames* e *Facebook*. Protótipos da interface gráfica da ferramenta *desktop* de validação de anotações foram feitos em colaboração com pesquisadores do Instituto Universidade Virtual, da Universidade Federal do Ceará [12]. Os próximos passos serão a implementação dos primeiros algoritmos de geração de anotações e testes de integração com o serviço de armazenamento de arquivos *Dropbox*.

7. COMPARAÇÃO COM TRABALHOS RELACIONADOS

A literatura apresenta várias ferramentas e estratégias para o suporte à anotação de fotos. A proposta apresentada se assemelha a trabalhos existentes [6] por fornecer funcionalidades de suporte à anotação tanto em dispositivos móveis como em ambiente *desktop*. Outra semelhança com soluções já implementadas [3][4] é a utilização de elementos do contexto do usuário como ponto de partida para o enriquecimento semântico das anotações geradas.

O diferencial da solução apresentada é a forma oportunista e transparente pela qual a geração de anotações é feita no serviço remoto que compõe a arquitetura. Assim, ao contrário de soluções existentes, algoritmos mais sofisticados e que demandam mais recursos podem ser executados em *background* enquanto o usuário captura novas fotos. As anotações geradas em *background* podem, a critério do usuário, ser utilizadas para enriquecer ainda mais suas fotos.

8. AGRADECIMENTOS

Este trabalho é parcialmente financiado pelo projeto FUNCAP de número PJP-0072-00091.01.00/12.

9. REFERÊNCIAS

[1] Chai, Y., Xia, T., Zhu, J. e Li, H. (2010). Intelligent digital photo management system using Ontology and SWRL. *International Conference on Computational Intelligence and Security*, 18-22.

- [2] Ma, H., Zhu, J., Lyu, M.R.-T. e King, I. (2010). Bridging the Semantic Gap Between Image Contents and Tags. *IEEE Transactions on Multimedia*, 462-473.
- [3] Viana, W., Filho, J. B., Gensel, J., Villanova-Oliver, M., Martin, H. (2007). PhotoMap—Automatic Spatiotemporal Annotation for Mobile Photos. *Web and Wireless Geographical Information Systems*, 187-201.
- [4] Xia, S., Gong, X., Wang, W., Tian, Y., Yang, X., Ma, J. (2010). Context-Aware Image Annotation and Retrieval on Mobile Device. *Second International Conference on Multimedia and Information Technology*, 111-114.
- [5] Naaman, M., Harada, S., Wang, Q., Garcia-Molina, H., Paepcke, A. (2004). Context data in georeferenced digital photo collections. *Proceedings of the 12th annual ACM international conference on Multimedia*, 196-203.
- [6] Viana, W., Miron, A. D., Moisuc, B., Gensel, J., Villanova-Oliver, M., Martin, H. (2010). Towards the semantic and context-aware management of mobile multimedia. *Multimedia Tools and Applications*, vol. 53, 391-429.
- [7] Dey, A. K., Abowd G. D. (1999). Towards a Better Understanding of Context and Context-Awareness. *Proceedings of the 1st international symposium on Handheld and Ubiquitous Computing*, 304-307.
- [8] Viana, W., Filho, J. B., Gensel, J., Villanova-Oliver, M., Martin, H. (2008). PhotoMap: from location and time to context-aware photo annotations. *Journal of Location Based Services*, vol. 3, no. 3, 211-235.
- [9] Figueirêdo, H. F., Lacerda, Y. A., Paiva, A. C., Casanova, M. A., Baptista, C. S. (2012). PhotoGeo: a photo digital library with spatial-temporal support and self-annotation. *Multimedia Tools and Applications*, vol. 59, no. 1, 279-305.
- [10] Rost, M., Cramer, H., Holmquist, L. E. (2012). Mobile exploration of geotagged photographs. *Personal and Ubiquitous Computing*, vol. 16, no. 6, 665-676.
- [11] Portugal, P. P. (2011). Mobile ContextPicasa: uma aplicação móvel para anotação contextual automática de fotos. Dissertação de Mestrado, Universidade Federal do Ceará.
- [12] Ribeiro, T. G., Novais, C. E. B. (2011). Prática do Design de Interação em uma experiência conjunta entre as disciplinas de IHC, Design de Interfaces e Semiótica. *Workshop sobre Ensino de IHC*.