

InTag: uma abordagem para Marcação e Inferência OWL em Imagens Georreferenciadas usando a Plataforma AEROGRAF

Tiago Josué Diedrich, Diego Geraldo, Henrique Costa Marques, José Maria Parente de Oliveira
Instituto Tecnológico de Aeronáutica (ITA) – Avenida Marechal Eduardo Gomes, S/Nº – DCTA – São José dos Campos – SP

Resumo – Para militares especializados, reconhecer e catalogar uma cena é uma tarefa importante. Por vezes, analistas não conseguem identificar todos os elementos contidos numa imagem, ou têm diferentes percepções daquilo que observam numa cena. Sabe-se, por exemplo, que o armazenamento de informações, por contexto semântico, permite explorar elementos de inteligência em imagens. Porém, os instrumentos que permitem esse feito ainda são limitados, devido a um processo de implementação restrito. Consequentemente, são necessários meios para realizar uma catalogação adequada, como forma de permitir identificação dos elementos ocultos. Assim, o artigo apresenta a elaboração de um *plugin* implementado na Plataforma AEROGRAF, para inserir marcação em imagens. As informações são convertidas para modelagem ontológica e, por meio do processo de inferência, oferecido pelo *software* Protégé, são evidenciados elementos não observados. Como estudo de caso aborda-se o domínio operacional de aeroportos de interesse da Força Aérea, para amparar um sistema de apoio à decisão.

Palavras-Chave – Engenharia de Ontologia, *Plugin*, Imagem.

I. INTRODUÇÃO

Atualmente (2014), poucos trabalhos publicados voltam-se para o assunto de Inteligência de Imagens. Da mesma forma, são poucas as pesquisas que destacam o uso de ferramentas implementadas, associadas à representação de conhecimento, ontologias e informações em imagens. Contudo, o interesse nessa área tem aumentado consideravelmente, devido ao processo de identificação de elementos imagéticos que representam uma cena [1].

A atividade de marcar, rotular, etiquetar ou inserir tags em imagens que representam cenas militares, por exemplo, refletem nos seguintes empregos: acompanhamento de movimentação de tropas, identificação de elementos de interesse, proteção de pontos de defesa amiga, observação e análise dos danos de batalha (*Battle Damage Assessment* – BDA), entre outros [2], porém para identificar os elementos que compõem uma cena de interesse militar é necessária a expertise de analistas capacitados. Observa-se que na Força Aérea Brasileira (FAB) existe pessoal especializado e dedicado para essa finalidade. São militares treinados para interpretar e marcar informações nas imagens. Assim, o trabalho desses interpretes volta-se para o processo de contextualização dos elementos contidos numa dada cena.

Contudo, analistas que têm por missão reconhecer alvos militares, por vezes, apresentam diferentes percepções daquilo que observam numa imagem. Mesmo realizando uma catalogação assertiva e eficiente, podem colaborar para um processo desarticulado da informação. O conhecimento inserido numa cena também pode se tornar irrelevante ou confuso, resultando em subjetividade, imprecisão, falta de

associação semântica e, eventualmente, pode apontar para áreas não desejadas [3].

Essa é uma situação típica nos Esquadrões de Reconhecimento da FAB. Essas Unidades Aéreas empregam Estações de Planejamento e Interpretação que, por sua vez, armazenam Banco de Dados de Imagens (BDIs). Por meio desses BDIs, uma grande quantidade de cenas de alvos militares são analisadas por meio da identificação assertiva dos elementos contidos numa imagem específica.

Ainda que a tarefa de marcar imagens seja limitada à interpretação subjetiva dos analistas, (elementos observados, situações evidenciadas, alvos militares identificados, atividades relacionadas, etc.), não existe um padrão perfeito de inserção de metadados em imagens, capaz de marcar as informações e inferir determinados conceitos [4].

Segundo [5], as pessoas que realizam interpretação de imagens nem sempre conseguem reconhecer alguns elementos importantes de uma cena, no momento da catalogação. Isso reduz a eficiência no processo da catalogação, e resulta em desperdício de recursos envolvidos [6].

Assim sendo, o problema abordado por este artigo diz respeito à omissão de informações inerentes à imagem no momento da catalogação, uma vez que nem todos os elementos que compõem a cena são observados, interpretados e marcados pelo analista [7].

Outro ponto refere-se ao processo de anotação em imagens. Essa tarefa muitas vezes ocorre sem uma padronização de termos descritores. Isso propicia problemas para identificar o contexto da mídia. Assim, para que seja possível melhorar a indexação desses termos, alguns autores propõem o uso de ferramentas que possibilitem interoperabilidade semântica, expressividade de conceitos ou inferências por meio de ontologias [8].

O presente trabalho introduz uma ferramenta capaz de indexar informações de inteligência a imagens por meio de descritores padronizados (metadados) e identificar conceitos ou informações não observadas por interpretes, no momento inicial da catalogação das imagens.

A implementação dessa ferramenta deu-se por meio do desenvolvimento de um novo *plugin* para a Plataforma AEROGRAF, um sistema de informações geográficas da FAB desenvolvido no Instituto de Estudos Avançados (IEAv). O *plugin*, chamado InTag, oferece uma interface adequada para anotação padronizada de imagens georreferenciadas e a exportação dos metadados em forma de ontologia. As informações são transferidas para o *software* Protégé [9], onde regras associadas ao domínio semântico são aplicadas aos dados originais, evidenciando novas marcações, por meio de inferência. Dessa forma é possível representar computacionalmente uma imagem e identificar elementos que muitas vezes não são observados por analistas.

Como estudo de caso, adotou-se o domínio de possíveis aeroportos de interesse militar. Assim, alguns elementos que representam esse contexto passam por um processo de inferência e identificação de marcações pouco evidentes.

II. REVISÃO BIBLIOGRÁFICA

Segundo [6], a ausência de descritores padronizados para representar imagens ocasiona uma dificuldade no processo de busca, de acesso e de recuperação de recursos informacionais da cena. A autora aponta para as dificuldades na implantação do uso dos códigos e formatos de marcações, para acompanhamento e controle de qualidade dos produtos de catalogação, bem como o problema de divergências linguísticas existentes. Descreve, então, propostas de como construir modos mais eficientes para representar imagens, para torná-las disponíveis, acessíveis e recuperáveis. Isso tudo é viabilizado por meio da padronização das *tags*, as quais orientam a descrição da imagem. Contudo, a autora espera encontrar resultados mais satisfatórios, por meio de um esquema de descrição de imagens digitais, que contemple a lógica descritiva já definida na área para os recursos informacionais tradicionais.

Já [10] define que a marcação do conteúdo em uma imagem, com descritores conhecidos como *tags*, geralmente não oferece uma padronização, por não possuir relação com uma categoria específica. O autor parte para a criação de um modelo de marcação, de forma colaborativa, permitindo inserção por diversos usuários, que prevê certos padrões estáveis, e relaciona-os na imagem. Ele analisa a estrutura de colaboração do sistema de marcação, bem como os seus aspectos dinâmicos, descobrindo os tipos de *tags* geralmente utilizadas. Isso permite organizar conceitos relevantes para recuperar informações futuras, por meio de navegação, de filtragem ou de pesquisa. A padronização de termos demonstra que marcações podem ser valiosas em conjunto, tanto quanto individualmente.

Conforme definido por [11], a marcação automática em imagens é uma atividade demandante e passível de erros, pois permite inconsistência de termos e/ou ausência de semântica. Muitas vezes ocorre a falta de exemplos de aprendizagem para modelar a diversidade visual que uma marca pode retratar. Além disso, muitas informações são ambíguas, limitadas em termos de completude e, excessivamente personalizadas. O autor procurou resolver o problema da seguinte forma: combinou marcação manual e automática em uma estrutura baseada em pesquisa; para uma imagem sem rótulo, recuperou a marcação por meio de seus vizinhos visuais, onde um usuário havia previamente catalogado, por meio de uma grande base de dados; selecionou marcações relevantes, para anotar imagens sem rótulos; e, para combater a insegurança e dispersão de marcação do usuário, foi introduzida uma modalidade para observar a relevância da etiqueta, que aborda uma forma textual eficiente por meio de dicas. Assim, o autor aproveitou as *tags* inseridas pelo usuário, para conseguir melhor a marcação automática. Ao identificar um rótulo relevante, o método abordado fez com que o processo de marcação permitisse uma recuperação mais robusta e assertiva, porém a semântica ainda ficou incoerente. O autor propõe melhorar a técnica.

As análises de [8] informam que imagens existentes em sites sociais muitas vezes recebem marcações inadequadas, errôneas e incompletas, o que degrada fortemente a eficácia

de abordagem do processo de catalogação e identificação dos conceitos relacionados. O autor propôs o *Sparse Tag Patterns* (STP), um modelo esparso para padronizar marcações, com o objetivo de descobrir incoerências, bem como padrões de *tags* inseridas por usuários, traduzindo os resultados em representações do significado. Devido à importância da marcação de imagem, que se dá semanticamente por meio de palavras-chave relevantes, a ferramenta permitiu recuperar, indexar, organizar e compreender grande coleção de dados de imagem. Consentiu, ainda, que cenas sociais englobassem o mesmo conceito semântico, e que usuários pudessem complementar, mutuamente, algumas marcações. Contudo, o autor não definiu a forma de identificar os conceitos não observados na imagem.

Para finalizar [12] afirma que há uma necessidade emergente para criação de novas técnicas de catalogação, pois as existentes não possuem uma construção padronizada, que represente sua informação coerentemente. Através da análise das teorias, dos princípios, dos fundamentos e dos instrumentos de representação da marcações, o autor propõem a utilização de metadados, para categorizar os elementos descritos na cena. Destaca, ainda, as metodologias que orientam a construção padronizada de representações por contexto associado, fundamentado nos avanços tecnológicos existentes. O autor demonstra a integração estratégica entre tecnologias de informática e o tratamento descritivo da informação, nos processos de representação do conhecimento, facilitando assim a busca, a recuperação, o acesso, a difusão e a aquisição da informação em ambientes digitais.

Assim, para identificar conceitos não observados em imagens, os autores do presente artigo propõem a utilização de *plugin*, por meio da plataforma AEROGRAF. Como forma de relacionar descritores padronizados, por meio de metadados agregados às imagens, esses são convertidos em formado ontológicos e submetidos a um processo de inferência e correlação semântica. Por contexto, são identificadas algumas informações ocultas, devido às regras inerentes à modelagem ontológica previamente definida, conforme domínio adotado.

III. METODOLOGIA

Como forma de contextualizar o trabalho desenvolvido segue-se a metodologia empregada:

1. Primeiramente foi realizada uma consulta à bibliografia, Norma OTAN [13] e Manual de Implementação de Aeroportos [14], para reforçar estudos e definir descritores padronizados, a serem incorporados como metadados às imagens pertencentes ao domínio específico de aeroportos;
2. Os conceitos foram arranjados em uma tabela, com o objetivo de prover um panorama claro e organizado;
3. Para formalizar a estruturação semântica das cenas, foi empregada a Plataforma AEROGRAF, por meio do *plugin* InTag. A Plataforma possibilita inserir informações às cenas, por meio de um sistema de anotação em imagens [15];
4. As informações foram organizadas em um BDI, construído no próprio AEROGRAF com a finalidade de correlacionar os metadados entre as cenas, bem como validar a utilização do InTag; e
5. Em seguida, através de análise, interpretação e marcação de imagem realizada, os dados foram estruturados na

AEROGRAF e, em seguida, convertidos para formato *Web Ontology Language* (OWL). Utilizou-se, para esse fim, o *software* Protégé, um editor de ontologias e aquisição de conhecimento. Por meio dele é possível incluir classificadores dedutivos para validar modelos e inferir novas informações [9].

IV. MATERIAIS E MÉTODOS UTILIZADOS

Para gerar os resultados esperados foi necessária a utilização dos seguintes aplicativos computacionais:

Plataforma AEROGRAF

Sistema que opera conforme *framework*, o qual permite incorporações de módulos com diferentes funcionalidades, chamados de *plugins*. A Plataforma AEROGRAF foi criada por meio de um projeto desenvolvido no Instituto de Estudos Avançados (IEAv), na década de 90. É um aplicativo computacional capaz de atender diferentes necessidades da FAB, dentre elas o Planejamento de Missão Aérea. A plataforma possui bases cartográficas e imagens georreferenciadas [15].

Protégé – Versão 4.3

Aplicativo computacional que permite criar e editar ontologias em formato *Web Ontology Language*, além de realizar inferências sobre a mesma. Consente o uso de regras, que possibilitam a identificação de elementos não observados ou pouco evidentes nas imagens, por meio de conceitos lógicos e axiomas agregados [9] [16-17].

V. PROCEDIMENTOS E RESULTADOS

A. Definição do Domínio

Utilizou-se para o presente trabalho a padronização dos termos relacionados ao domínio. Para tanto, consultou-se a Norma *North Atlantic Treaty Organization* (NATO) *Standardization Agreement* (STANAG) 3596 [13].

Além da Norma STANAG, utilizou-se o Manual de Implementação de Aeroportos [14], que possibilita análise das características que compõe determinados tipos de aeroportos, os quais são definidos pelo extinto Departamento de Aviação Civil (DAC), mas que atendem perfeitamente para o propósito dessa pesquisa.

Essas publicações são ostensivas e permitiram as definições dos descritores padronizados, como forma de apresentar o domínio de imagens de aeroportos de interesse da FAB.

B. Desenvolvimento do Plugin InTag

O *plugin* InTag foi implementado com o auxílio da equipe de desenvolvimento da Plataforma AEROGRAF do IEAv. O aplicativo foi concebido para fornecer uma interface adequada para a marcação padronizada de imagens georreferenciadas.

Na Fig. 1 observa-se uma imagem parcial da Plataforma AEROGRAF, como forma de exibir o *plugin* criado.

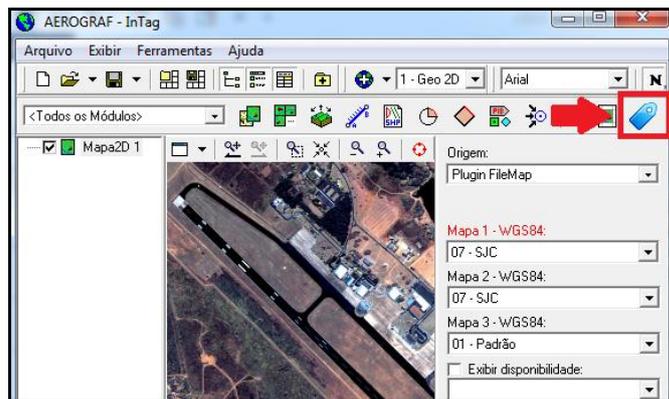


Fig. 1. Plugin InTag

Após a criação do *plugin* confeccionou-se um BDI, com o propósito de apresentar informações de algumas cenas, por meio de metadados. Foram selecionadas algumas imagens de aeroportos e inseridas informações observadas em cada localidade. Esse BDI serviu como suporte para apoiar o padrão de pesquisa frente à geração da modelagem ontológica.

O domínio específico, que representa possíveis aeroportos de interesse militar, são selecionados ao se criar uma instância do *plugin* InTag na Plataforma AEROGRAF, para definir os descritores padronizados, conforme Fig. 2.

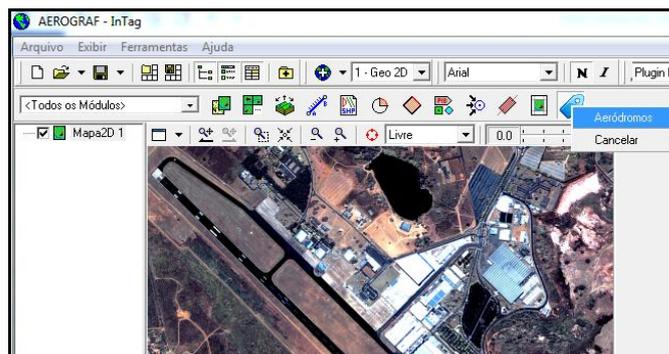


Fig. 2. Domínio previamente especificado: “Aeródromos”.

Para esse trabalho adotou-se o aeroporto de São José dos Campos.

As informações foram inseridas por meio de metadados, tratados com a utilização da Plataforma AEROGRAF. Como exemplo, observa-se a informação de “Terminal de passageiros” sendo inserida, conforme Fig.3.

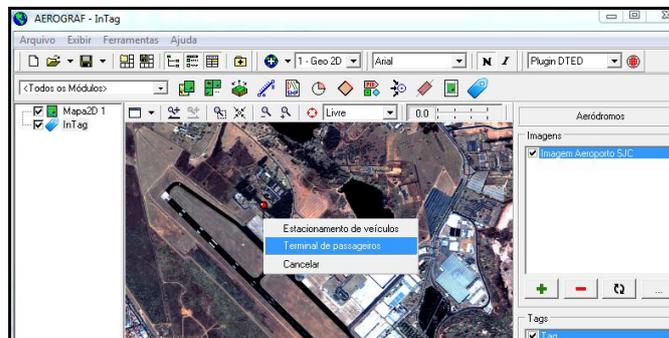


Fig. 3. Inserção de informações do descritor padronizado “Terminal de passageiros”, por meio de metadados.

Em seguida, iniciou-se a inserção de todas as informações observadas pelo interprete, que seguem: revestimento de

asfalto, aeródromo que opera por instrumento de precisão, indicador de direção do vento iluminado, sala de serviço de informação aeronáutica, estação permissionária de telecomunicação aeronáutica de categoria B, terminal de passageiros, estacionamento de veículos, equipamento indicador do ângulo de descida, radio farol não direcional, radio farol omnidirecional, balizamento noturno, farol rotativo de aeródromo, serviço de salvamento de combate contra incêndio, torre de controle, sistema de pouso por instrumento e sistema de luzes de aproximação.

Todos os elementos foram inseridos na imagem, conforme experiência e interpretação do analista, no momento da catalogação. Assim, cada informação inserida, visualiza-se com o uso da simbologia de um alfinete.

Para ler os metadados vinculado a um marcador específico, por exemplo, “Revestimento de asfalto”, o descritor apresenta em “caixa” seu Domínio, Classe e Tag inserida, como é o caso da informação evidenciada na Fig. 4.

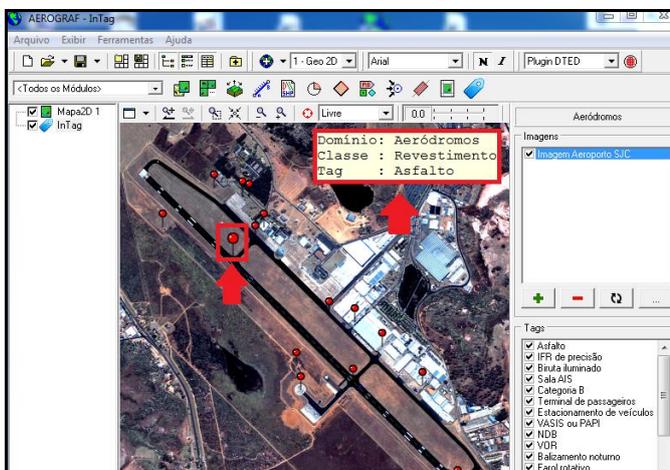


Fig. 4. Metadados inseridos na imagem.
Tag “Revestimento de asfalto” em evidência.

De acordo com as informações definidas pelo Manual de Implementação de Aeroportos [14], uma localidade que possui todos esses elementos dissertados anteriormente, obrigatoriamente enquadra-se na categoria de “Aeroporto de Aviação Domestica Nacional”. Infelizmente, na imagem analisada, o interprete não reconheceu e nem inseriu esse marcador.

Assim, para validar o experimento, bem como identificar os elementos pouco evidentes da cena, os metadados foram exportados para formato OWL, uma linguagem recomendada pela World Wide Web Consortium (W3C) [18], conforme Fig. 5.

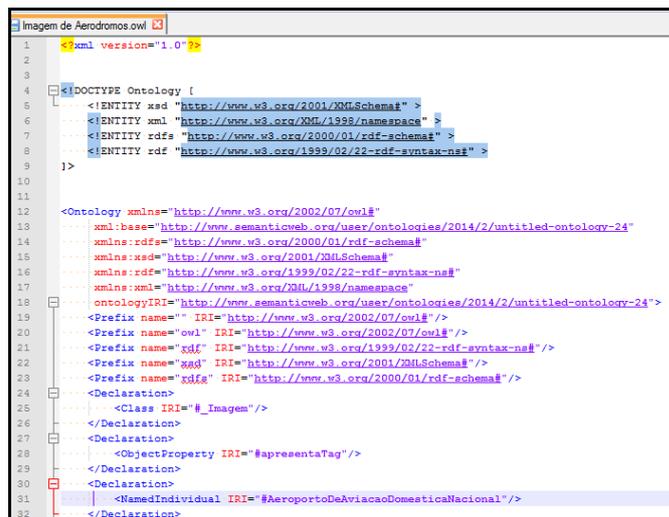


Fig. 5. Visualização parcial do arquivo OWL gerado a partir da Plataforma AEROGRAF.

C. Software Protégé

Nessa etapa foram utilizados os recursos oferecidos pelo software Protégé, versão 4.3 no padrão OWL [9], como forma de representar a ontologia para proporcionar interação entre o domínio de aeródromos e a modelagem oferecida pela AEROGRAF.

Assim, por meio do Protégé definiram-se regras. Uma dessas regras impõe que em todas as imagens que contém os elementos (metadados) apresentados no exemplo anterior (imagem do aeroporto de São José dos Campos) deve ser inserida uma nova marca. Nesse caso, reflete o tipo de categoria a que o aeroporto pertence: “Aeroporto de Aviação Domestica Nacional”. Observa-se a regra gerada na Fig. 6.

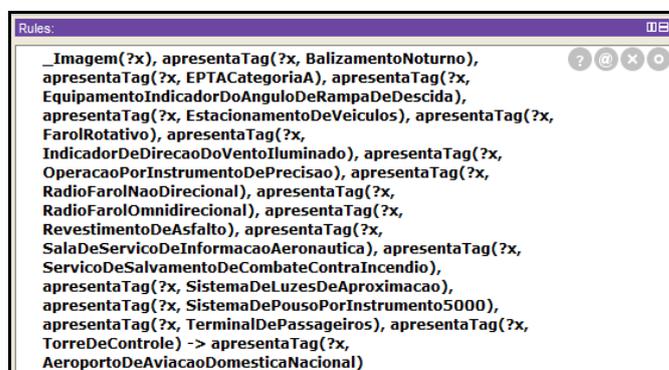


Fig. 6. Definição da Regra para “Aeroporto de Aviação Doméstica Nacional”.

Como referência, o Protégé foi utilizado por representar a semântica de relacionamentos de todas as entidades de domínio (anotações) dissertadas nas imagens, além de ser um aplicativo gratuito. É fundamentado em Java, e recomendado pela W3C [9][16-18].

Após a exportação das informações da imagem, o arquivo OWL foi inserido no Protégé, juntamente com a regra formulada. Foi acionado um mecanismo de inferência e identificada a informação oculta, conforme Fig. 7.

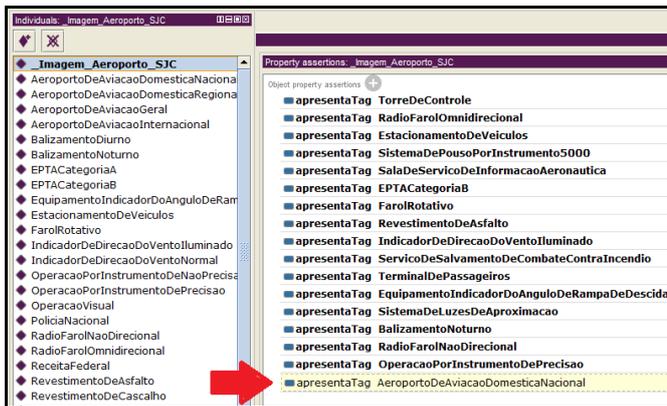


Fig. 7. Visualização da informação oculta através do software Protégé.

A informação da categoria de aeroporto, não identificada inicialmente pelo analista, foi descoberta devido à regra inferida pelo Protégé.

Assim, observa-se que a modelagem ontológica, confeccionada a partir da Plataforma AEROGRAF, juntamente com a regra explicitada, serviu como suporte para apoiar o padrão de pesquisa frente ao domínio de Aeroportos de interesse militar, identificando uma informação não observada pelo analista quando no momento da marcação da imagem.

VI. CONCLUSÕES E RECOMENDAÇÕES

A. Conclusões

O processo de definir descritores padronizados permitiu conceitualizar o domínio de imagens de aeroportos de interesse da FAB.

A utilização da Plataforma AEROGRAF, por meio do novo *plugin* InTag, possibilitou a inserção dos descritores padronizados de forma coerente e assertiva, permitindo o uso do processo de inferência.

Através desse trabalho foi possível verificar que a criação do *plugin* permite ampliar outros domínios de interesse. Assim, caso o usuário defina previamente descritores padronizados para um domínio específico diferente, um conjunto de regras pode identificar elementos ocultos em qualquer mídia.

A modelagem em formato OWL, representando o padrão definido pela W3C, permite integrar metadados a novas linguagens da WEB.

O software Protégé possibilitou o processo de inferência de maneira satisfatória e completa, pois identificou assertivamente a anotação não observada pelo analista.

A representação do conhecimento, por meio da conceitualização oferecida pela AEROGRAF, apresentou novos conceitos relativos à imagem, expondo uma nova informação não observada inicialmente.

Pretende-se incorporar a criação de regras e a fase de inferência às capacidades do *plugin* InTag.

Como resultado, o uso do *plugin* proposto demonstra a capacidade de transformar conceitos padronizados em metadados de imagens. Esses são submetidos a uma ontologia de domínio, a qual viabiliza inferência contextualizada. A questão ontológica adiciona novas informações inferidas, as quais apontam para os elementos não evidenciados durante o processo de catalogação. Isso

propicia uma melhoria no processo de tomada de decisão pelos elos de Comando e Controle (C2).

B. Recomendações

Recomenda-se ampliação do número de imagens, como forma de analisar a complexidade de amplos BDIs.

Propõe-se o uso de novas regras, para estender a OWL de maneira coerente, com a finalidade de integrar amplamente o contexto semântico agregado às imagens.

Sugere-se a confecção de perguntas que possam recuperar informações de interesse estratégico, conforme a necessidade do usuário, com o fim de oferecer mais respaldo para a tomada de decisão dentro dos conceitos de C2.

REFERÊNCIAS

- [1] T. Schreiber, B. Dubbeldam, J. Wielemaker, and B. Wielinga, "Ontology-based photo annotation," *Intelligent Systems, IEEE*, vol. 16, pp. 66-74, 2001.
- [2] Força Aérea Brasileira, "DCA 1-1 Doutrina Básica da Força Aérea Brasileira", 2012, Brasil.
- [3] S. Ghosh and S. K. Bandyopadhyay, "A Tutorial Review of Automatic Image Tagging Technique Using Text Mining," *International Journal of Research in Engineering and Technology*, vol. 2, pp. 282 - 289, 2013.
- [4] A. C. Simionato, "Representação, Acesso, Uso e Reuso da Imagem Digital," Mestrado em Ciência da Informação, Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, 2012.
- [5] L. Jia and J. Z. Wang, "Real-Time Computerized Annotation of Pictures," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, v. 30, n. 6, pp. 985-1002, 2008.
- [6] A. C. Simionato, "Catalogação de Imagens Digitais," vol. 9, ed: Revista de Iniciação Científica da FFC, pp. 116-129, 2009.
- [7] S. Gao, Z. Wang, L. Chia, I. W. Tsang, "Automatic image tagging via category label and web data", School of Computer Engineering, Nanyang Technological University, Singapore. ACM Special Interest Group on Multimedia, pp.1115-1118, 2010.
- [8] J. Lin, J. Yuan, L.-Y. Duan, S. Luo, and W. Gao, "Social Image Tagging by Mining Sparse Tag Patterns from Auxiliary Data," in *Multimedia and Expo (ICME), IEEE International Conference*, pp. 7-12, 2012.
- [9] Software Protégé, "Semantic Technology & Business Conference," 2014. Disponível em: <<http://protege.stanford.edu/>>
- [10] S. A. Golder and B. A. Huberman, "Usage patterns of collaborative tagging systems," *J. Inf. Sci.*, vol. 32, pp. 198-208, 2006.
- [11] X. Li, C. G. M. Snoek, and M. Worring, "Annotating images by harnessing worldwide user-tagged photos," *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2009.
- [12] R. C. V. Alves, "Metadados como Elementos do Processo de Catalogação," Doutorado em Ciência da Informação, Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010.
- [13] NATO, "Standardization Agreement - STANAG 3596 - Air Reconnaissance Requesting and Target Reporting Guide," 5 ed, 2003.
- [14] BRASIL, Comando da Aeronáutica, Instituto de Aviação Civil, "Manual de Implementação de Aeroportos", 65p, 2004. Disponível em: <<http://www.anac.gov.br/arquivos/pdf/manualImplementacaoGeral.pdf>>
- [15] D. Geraldo, "Trajetórias táticas para veículos aéreos não tripulados : uma abordagem por otimização multiobjetivo e evolução diferencial," Mestrado em Engenharia Aeronáutica e Mecânica, Instituto Tecnológico de Aeronáutica, São José dos Campos - SP, 2012.
- [16] M. K. Smith, C. Welty, and D. L. McGuinness, "OWL Web Ontology Language Guide," ed: W3C Recommendation Service, 2004.
- [17] M. Horridge, H. Knublauch, A. Rector, R. Stevens, C. Wroe, S. Jupp, et al. (2011) A Practical Guide To Building OWL Ontologies Using Protégé 4 and CO-ODE Tools. Disponível em: <http://130.88.198.11/tutorials/protegeowltutorial/resources/ProtegeOWLTutorialP4_v1_3.pdf>.
- [18] W3C Recommendation Service. OWL Web Ontology Language Guide. Disponível em: <<http://www.w3.org/TR/owl-guide/>>.