

# Reconhecimento de Objetos Baseado em Contexto Utilizando a Lógica de Descrição Probabilística *CRA $\mathcal{L}\mathcal{C}$*

Valquiria Fenelon Pereira<sup>1</sup>, Paulo Eduardo Santos<sup>2</sup>, Fabio Gagliardi Cozman<sup>1</sup>

<sup>1</sup>Escola Politécnica — Universidade de São Paulo  
São Paulo – SP – Brasil

<sup>2</sup>Centro Universitário da FEI  
São Bernardo do Campo – SP – Brasil.

{vfenelon, fgcozman}@usp.br, pesantos@fei.edu.br

**Abstract.** *Robotic agents still lack the ability to reason about objects in real scenes. We propose an approach in which the context of objects is modeled using a probabilistic description logic (CRA $\mathcal{L}\mathcal{C}$ ), and the data extracted by computer vision methods provide information for inferences concerning objects. In our experiments the SURF algorithm (Speed Up Robust Features) was used to obtain object descriptors in images captured from two different environments. The results show an increase in positive responses due to context modeling with CRA $\mathcal{L}\mathcal{C}$ .*

**Resumo.** *Agentes robóticos ainda não têm a habilidade de raciocinar sobre objetos em cenas reais. Neste artigo propomos uma abordagem em que o contexto de objetos é modelado por meio de uma lógica de descrição probabilística (CRA $\mathcal{L}\mathcal{C}$ ), e onde os dados extraídos por visão computacional são fontes de informação para inferências sobre objetos. Em nossos experimentos usamos o algoritmo SURF (Speed Up Robust Features) para obter descritores de objetos em imagens capturadas de dois ambientes distintos. Os resultados apresentam um aumento nas respostas positivas devido à modelagem de contexto por meio de CRA $\mathcal{L}\mathcal{C}$ .*

## 1. Introdução

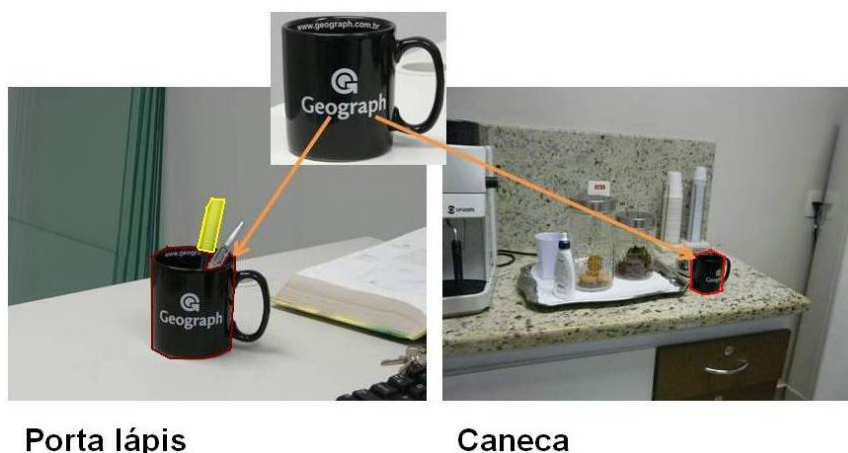
Na interação de um robô com o ambiente, os objetos estão sempre em evidência. Por exemplo, uma das capacidades cruciais, para que um robô móvel consiga interagir e executar tarefas domésticas, é perceber os objetos e seus significados em ambientes não restritos [Ramisa et al. 2011]. Através de imagens capturadas e aplicando algoritmos de visão computacional, podemos obter informações úteis sobre os objetos que estão em um mundo físico real [Shapiro and Stockman 2001]. Em ambiente real, um agente trabalha com incertezas; o reconhecimento de objetos torna-se mais complicado em ambientes complexos, nos quais os objetos podem fazer parte de um arranjo espacial não trivial. Estes ambientes são um desafio, deste modo todas as informações deve ser utilizadas [Ramisa et al. 2011].

Vários estudos exploram o contexto de objetos para desambiguar imagens [Parizi et al. 2009, Semenovich and Sowmya 2009]. A maior parte desses estudos ainda explora as imagens a partir de bancos de dados, algumas direcionadas a algoritmos de busca [Okabe et al. 2010] enquanto outras se relacionam com robôs móveis

[Ribes et al. 2008]. Em [Rabinovich 2007] e [Okabe et al. 2010] foi utilizado um modelo das relações entre os objetos em contexto em que foi estimada a probabilidade de objetos ocorrerem juntos em uma mesma cena.

Neste trabalho, para raciocinar sobre objetos, contexto e lidar com incerteza, utilizamos a lógica *Credal ALCC* (CRALCC) [Cozman and Polastro 2008], que é uma extensão probabilística da lógica de descrição *ALCC* [Schmidt-Schauss and Smolka 1991]. Podemos encontrar uma proposta similar de aplicação dessa lógica em [Santos et al. 2010, Fenelon et al. 2010], onde em um domínio de tráfego a localização do carro e o sentido das faixas de direção foram inferidas a partir do conhecimento a priori do contexto e dos dados dos sensores (GPS, mapa digital e câmera de vídeo). Em nosso experimento usamos o algoritmo SURF (*Speed Up Robust Features*) [Bay et al. 2006] para obter descritores de objetos em imagens capturadas de dois ambientes distintos. A identificação dos objetos leva então em conta os descritores e informações sobre contexto (Figura 1).

O artigo está organizado do seguinte modo: na próxima seção apresentaremos elementos básicos como o algoritmo de visão computacional SURF [Bay et al. 2008], uma breve revisão sobre modelagem de contextos [Bar 2004], *Credal ALCC* [Cozman and Polastro 2008]. Na Seção 3 discute-se a modelagem do contexto dos ambientes usando CRALCC. Os experimentos e resultados são descritos na Seção 4.



**Figura 1. Um objeto que apresenta uma ambiguidade quanto a sua função, que pode servir tanto como caneco ou porta-lápis dependendo do contexto e da relação entre os objetos do contexto.**

## 2. Elementos Básicos

A proposta desse artigo se baseia em definir o ambiente a partir de seus objetos típicos: ou seja, selecionamos de cada ambiente um conjunto de objetos que formam um *contexto*. Cada objeto tem um vetor descritor que o identifica, deste modo focamos na probabilidade de ser um contexto  $C$  dado os descritores do conjunto de objetos  $D$ :

$$\text{contexto}_i = \operatorname{argmax} P(C|D). \quad (1)$$

Para identificar um objeto que assume diferentes funções em diferentes contextos, optamos por primeiro reconhecer os descritores de objetos que formam o contexto

como do objeto em questão. Nesse processo conjunto estamos interessados em melhorar a identificação mesmo com informações parciais ou incompletas. Os algoritmos *Scale Invariant Feature Transform* (SIFT) [Lowe 2004] e *Speeded Up Robust Features* (SURF) [Bay et al. 2008] reconhecem objetos em imagens sem segmentação e tem uma tolerância a variação de tamanho. Ambos os algoritmos tem um detector de pontos chaves e descritores. O objeto é localizado comparando seus descritores com aqueles encontrados na imagem [Bauer et al. 2007]. Para obter descritores, o algoritmo de reconhecimento SURF [Bay et al. 2006, Bay et al. 2008] foi o escolhido, por estar disponível em bibliotecas de processamento de imagem *OpenCV* [Bradski 2000]. A próxima subseção apresenta esse algoritmo.

### **2.1. SURF - *Speed Up Robust Features***

O SURF é um detector de características invariáveis em escala e rotação em imagens [Bay et al. 2008, Bay et al. 2006]. Seu objetivo principal é encontrar correspondências entre duas imagens [Bay et al. 2006]. A busca por correspondências entre imagens pode ser dividida em três partes:

1. Encontrar pontos chaves na imagem. Definindo pontos de interesse como aqueles que sejam encontrados em diferentes condições de imagem.
2. A vizinhança dos pontos de interesse é representada por um vetor de características ou descritores, os quais devem ser distintivos e robustos ao ruído, às deformações geométricas da imagem e aos erros.
3. Comparar os vetores descritores das imagens. Essa comparação é geralmente baseada na distância entre vetores.

No algoritmo SURF os pontos chaves são localizados utilizando um detector Hessiano [Bay et al. 2006, Lindeberg 1998]. Para determinar orientação é utilizada a transformada de Haar (*Haar wavelets*) [Lee and 1999] e os descritores são formados a partir dos pulsos dos pontos ao redor do ponto chave. O vetor descritor tem 64 elementos, podendo ser dobrado se necessário. Note-se que a dimensão do descritor atua diretamente no tempo de execução da tarefa, portanto para maior agilidade computacional menores dimensões de vetores são desejáveis, todavia vetores menores são menos distintivos que os maiores [Bay et al. 2008], mesmo assim segundo [Bauer et al. 2007], dobrar o vetor não significa uma melhora nos resultados.

### **2.2. Contexto**

Um contexto pode ser definido tanto pela forma de relação espacial entre os objetos dentro de um ambiente, quanto pela provável ocorrência dos objetos num mesmo cenário [Semenovich and Sowmya 2009]. Experimentos com pessoas mostram que a organização espacial entre dois objetos tem um efeito significativo no reconhecimento dos mesmos, e o reconhecimento de um objeto ambíguo é mais efetivo quando este está associado a outro objeto [Bar et al. 1996, Bar 2004]. Desta forma, o uso de contexto pode ajudar na tarefa de reconhecimento de objetos quando as características intrínsecas do objeto não são claras o suficiente. Em geral cada ambiente tem um conjunto de objetos típicos, assim neste trabalho o contexto será definido somente pela ocorrência desses objetos em cena, ou seja, sem considerar as relações espaciais entre eles.

O algoritmo SURF tem bom desempenho para localizar objetos específicos com muitos detalhes, mas não é tão eficaz quando o objeto tem menos pontos de

destaque. O contexto descrito com a lógica probabilística serve para complementar este e outros algoritmos de reconhecimento de objetos [Parizi et al. 2009, Bohg et al. 2009, Jain et al. 2010]. Informação contextual será modelada na lógica de descrição probabilística descrita na próxima seção.

### 2.3. Credal $\mathcal{ALC}$

A lógica *Credal ALC* ( $CRALC$ ) [Cozman and Polastro 2008, Polastro and Cozman 2008] é uma extensão probabilística da lógica de descrição  $\mathcal{ALC}$  [Schmidt-Schauss and Smolka 1991]. O vocabulário básico da lógica de descrição  $\mathcal{ALC}$ , contém indivíduos, conceitos (conjunto de indivíduos) e papéis (relações binárias entre indivíduos) [Baader et al. 2002]. Além desses recursos, a versão probabilística aqui adotada,  $CRALC$ , assume que do lado esquerdo das inclusões ou definições aparecem somente conceitos. Para garantir a aciclicidade, não é permitido a nenhum conceito utilizar a si mesmo em inclusões ou definições determinísticas (ou probabilísticas), e que, dado um nome de conceito  $A$ , e um conceito  $B$  e um papel  $r$ , as seguintes asserções probabilísticas são possíveis:

$$P(A) \in [\underline{\alpha}_1, \bar{\alpha}_1], \quad (2)$$

$$P(A|B) \in [\underline{\alpha}_2, \bar{\alpha}_2], \quad (3)$$

$$P(r) \in [\underline{\beta}, \bar{\beta}]. \quad (4)$$

Os valores de probabilidade  $\alpha_i$  e  $\beta$  estão escritos como intervalos de probabilidades. Deste modo escrevemos  $P(A|B) = \alpha$  com um único valor quando  $\underline{\alpha} = \bar{\alpha}$ ,  $P(A|B) \geq \underline{\alpha}$  quando a maior probabilidade for igual a um e  $\underline{\alpha} < \bar{\alpha} = 1$ , e seguindo a  $P(A|B) \leq \bar{\alpha}$  para  $\underline{\alpha} = 0 < \bar{\alpha}$ .

A interpretação de indivíduos, conceitos e papéis, bem como dos operadores de  $\mathcal{ALC}$ , segue exatamente a semântica de  $\mathcal{ALC}$ . Adicionalmente, a Expressão (2) é interpretada como segue: para qualquer individuo  $x$  que pertença a um domínio  $\mathcal{D}$ , a probabilidade que  $x$  pertença a interpretação  $\mathcal{I}$  de  $A$  está no intervalo  $[\underline{\alpha}, \bar{\alpha}]$ , ou seja:

$$\forall x \in \mathcal{D} : P\left(\left\{\mathcal{I} : x \in \mathcal{I}(A)\right\}\right) \in [\underline{\alpha}, \bar{\alpha}].$$

Informalmente, podemos representar essa semântica como:

$$\forall x \in \mathcal{D} : P(A(x)) \in [\underline{\alpha}, \bar{\alpha}].$$

De forma similar, as semânticas das Expressões (3) e (4) são:

$$\begin{aligned} \forall x \in \mathcal{D} : P(A(x)|B(x)) &\in [\underline{\alpha}, \bar{\alpha}], \\ \forall (x, y) \in \mathcal{D} \times \mathcal{D} : P(r(x, y)) &\in [\underline{\beta}, \bar{\beta}]. \end{aligned}$$

Dado um domínio finito, um conjunto de sentenças em  $CRALC$  especifica probabilidades sobre todos os conceitos e papéis instanciados. As seguintes suposições garantem que uma única distribuição de probabilidades é especificada por um conjunto de sentenças: nome-único, probabilidade-pontual na avaliação, rigidez dos nomes, bem como uma condição de Markov que impõe independências acíclicas entre asserções



Figura 2. Ambiente do (a) laboratório e (b) da copa utilizados.

[Cozman and Polastro 2008]. Com estas suposições, um domínio finito e um conjunto de sentenças especificam uma única rede Bayesiana [Korb and Nicholson 2003] sobre todos os conceitos e papéis instanciados. Para obter a probabilidade uma asserção em  $CRALC$  gera uma rede Bayesiana a partir do conjunto de sentenças, e realiza inferência nesta rede. Assim uma definição  $A \equiv C$  é traduzida para  $P(A|C) = 1$  e  $P(A|\neg C) = 0$ , ou seja, se  $A$  e  $C$  são idênticos logo, dado  $C$  então  $A$ . A inclusão  $A \sqsubseteq C$  é representado pela  $P(A|C) = \alpha$  e  $P(A|\neg C) = 0$ , onde  $\alpha$  é um valor de probabilidade de  $A$  dado  $C$ .

### 3. Modelando Contexto com $CRALC$

O modelamento de contextos contém informações sobre os objetos, tais como a identidade e o posicionamento espacial [Bar 2004]. Ao escrever o contexto em  $CRALC$  algumas restrições devem ser observadas, tal como aciclicidade. Usamos como base para nossa modelagem a ontologia de contexto de ambientes apresentada em [Polastro et al. 2010] feita para localização de um robô móvel (note-se que a ontologia em [Polastro et al. 2010] foi utilizada assumindo informações sobre objetos, o que não assumimos nesse trabalho).

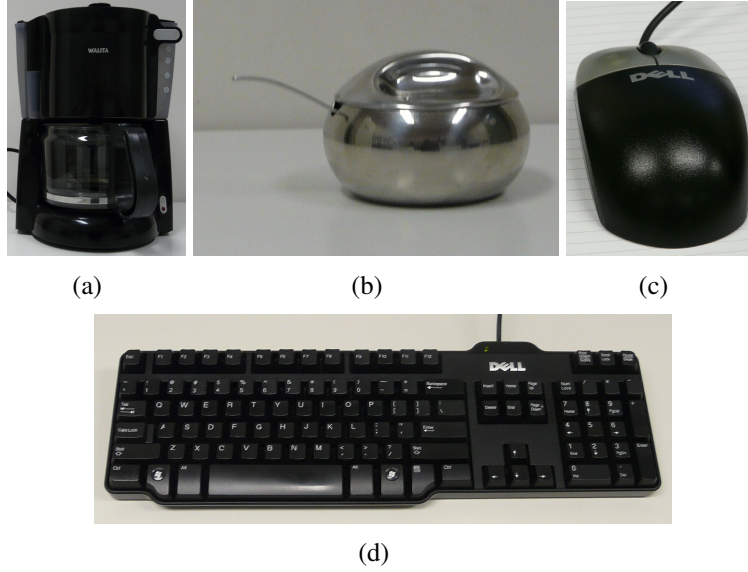
O ponto de partida foi uma tarefa concreta: rotular um objeto como copo ou porta-lápis (Figura 1). A identidade do objeto muda dependendo do ambiente que se encontra. As Expressões (5) e (6) mostram o conceito para nomear o objeto como copo ou porta-lápis respectivamente.

$$Copo \equiv Objeto \sqcap Cozinha \quad (5)$$

$$Porta - lápis \equiv Objeto \sqcap Laboratório \quad (6)$$

Os ambientes copa e do laboratório (Figura 2) foram então modelados com uma base de conhecimento definindo o contexto. Os contextos *Cozinha* e *laboratório* são conceitos que contém determinados objetos (Expressões (7) e (8)). Os objetos terão seus descritores armazenados em base de dados que serão obtidos de um banco de modelos. Os descritores localizados na imagem do ambiente com SURF são comparados com os descritores da base de dados. O banco de modelos é formado por imagens de objetos típicos de cada ambiente. Os contextos são naturalmente (Expressão (9)), já que um agente não pode estar fisicamente em dois ambientes distintos ao mesmo tempo:

$$Laboratório \sqsubseteq \exists contain.Li \quad (7)$$



**Figura 3. Alguns objetos típicos da copa utilizada e do laboratório utilizados como modelos. (a) Cafeteira, (b) açucareiro, (c) mouse e (d) teclado.**

$$Cozinha \sqsubseteq \exists \text{contain}.Ci \quad (8)$$

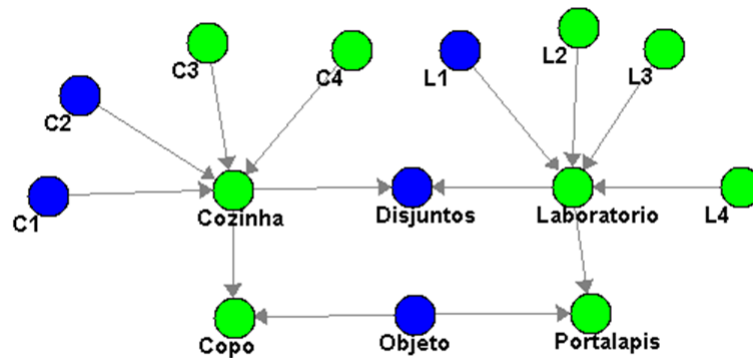
$$Cozinha \equiv \neg \text{Laboratório} \quad (9)$$

Os objetos modelos são conceitos primitivos. A probabilidade de ser um determinado contexto aumenta quanto mais objetos modelos típicos são encontrados neste ambiente. Foi definido um conjunto  $Ci$  que define os objetos típicos da cozinha, enquanto um conjunto  $Li$  define os objetos típicos do laboratório, onde  $i = 1, \dots, n$ . Na Figura 3 vemos alguns exemplos de objetos que aparecem em nossos ambientes e foram utilizados como modelos para definir o contexto.

$$\begin{aligned} P(Ci) &= \alpha_i \\ P(Li) &= \beta_i \\ P(\text{objeto}) &= \gamma \end{aligned}$$

Tanto  $\alpha_i$ ,  $\beta_i$  e  $\gamma$  são valores de probabilidade; conforme usado em [Polastro et al. 2010], foi escolhido o valor neutro 0,5 que não traz nenhuma informação a princípio. Os valores de  $\alpha_i$ ,  $\beta_i$  e  $\gamma$  podem ser alterados a partir dos dados coletados do sistema de visão. O valor de probabilidade será igual a 1 quando da evidência do objeto presente em cena, ou terá valor diferente considerando a quantidade de descritores do objeto em cena, assim quanto mais descritores localizados maior o valor de probabilidade.

Este domínio foi representado pela rede Bayesiana da Figura 4, que é o resultado da formalização em  $CRA\mathcal{L}C$ . Nesta rede a probabilidade a priori dos objetos será modificada pelos dados do sistema de visão. O vetor descritor das imagens dos ambientes é comparado com os vetores dos objetos a ser reconhecido e com o vetor dos objetos do contexto. Para cada imagem do ambiente e cada objeto forma-se então um vetor com os pares dos descritores semelhantes entre ambos. Desse modo extraímos as informações para formar as probabilidades de cada objeto.



**Figura 4. Rede Bayesiana que determina se objeto é rotulado como copo ou porta-lápis considerando o contexto.**

Os dados podem ser utilizados de dois modos: no primeiro, com a resposta positiva do objeto e dos objetos modelos localizados altera-se a probabilidade destes para um (1), independente do número de pares encontrados; no segundo, com a resposta do objeto e dos objetos modelos dependendo do tamanho do vetor com pares de cada cena, esse valor será utilizado como evidência em um intervalo de valor de probabilidade  $[0.5, 1]$ . Em ambos os casos quando o objeto não é localizado o valor de probabilidade a priori não se altera (0.5). Ou seja, somente os casos afirmativos são alterados na rede Bayesiana [Korb and Nicholson 2003], todos os outros conceitos primitivos permanecem com uma probabilidade neutra.

#### 4. Experimento

O experimento tem como objetivo investigar se os descritores invariantes de algoritmos de reconhecimento poderiam ser utilizados como informações, e como usar o contexto formalizado em lógica probabilística para tratar com as incertezas. A tarefa consiste em localizar um objeto que apresenta as mesmas características, porém com funções distintas dependendo do ambiente e das relações espaciais com os objetos deste contexto. No caso escolhemos um objeto, por exemplo, o caneco da Figura 5(a) que ao aparecer na cozinha ele é um copo, e quando no laboratório ele serve de porta lápis.

Com a câmera de um robô móvel *ActiveMedia PeopleBot<sup>TM</sup>* foram capturadas 200 imagens dos dois ambientes, tanto do laboratório quanto da copa. Destas 100 imagens, 66 contém o objeto alvo. Os objetos característicos de cada ambiente aparecem em número variado (ou seja, pode aparecer de um a quatro objetos), oclusos parcialmente ou não. Não existe restrição quanto a cor, iluminação, nitidez. As imagens tem uma resolução de 640x480. As Figuras 5(b) e 5(c) são dois exemplos das imagens capturadas no laboratório e na copa respectivamente.

Concluindo, nesta tarefa o algoritmo funciona com os seguintes passos:

- Os descritores da cena serão comparados aos descritores dos objetos (objeto alvo e contexto). Um exemplo da localização dos pontos do objeto e a cena está representada na Figura 5(d), onde cada linha liga um descritor do objeto com o seu par na cena.



(a)

(b)

(c)



(d)

**Figura 5. (a) Objeto alvo utilizado no experimento. Este objeto foi inserido em dois contextos distintos e deve ser nomeado como copo ou porta-lápis dependendo do ambiente. (b) Imagem capturada do laboratório e (c) imagem capturada na copa nenhuma restrição ou preparo do ambiente. (d) Localização de pontos de objeto e cena.**



- Quantificar os objetos com mais descritores encontrados e, a partir desses dados, definir o ambiente [Cozman and Polastro 2008].
- Rotular objetos ambíguos.

Na Tabela 1 a primeira coluna representa o modelo que deve ser reconhecido na imagem (ambiente, objeto). A segunda coluna apresenta o rótulo padrão escolhido para o objeto (copo ou porta-lápis). São 200 imagens sendo 100 imagens do laboratório e 100 imagens da copa. Existem 66 imagens com o objeto e 34 sem o objeto em cada ambiente. Nos experimentos utilizando somente o SURF para reconhecimento do objeto por uma de suas funções, obtivemos 17% de acurácia quando o objeto copo seria o localizado, e 29% de acurácia quando o porta-lápis deveria ser reconhecido (podemos ver esses valores na segunda coluna). Como o contexto não foi utilizado, não houve o reconhecimento do ambiente. A coluna quatro contém os valores de acurácia com o modelo de contexto utilizando o reconhecimento de objeto como evidência. Para o sistema onde o copo era padrão o valor 31% e quando o porta-lápis era o padrão 40%, nota-se uma melhora no resultado com a entrada do sistema do contexto. A última coluna mostra a acurácia para os casos de reconhecimento do objeto com contexto e *CRA<sub>CC</sub>*, utilizando o número de descritores como um valor de probabilidade. A alteração no reconhecimento do ambiente passou a ser de 38% (copo) e 41% (porta-lápis), mostrando uma melhora significativa na acurácia.

A Tabela 1 mostra os resultados da acurácia do experimento de rotular um objeto conforme sua função. Ou seja, os valores desta tabela representam a proporção entre resultados verdadeiros (positivos e negativos) e todos os resultados. Resultados verdadeiros positivos são as cenas com o objeto rotulado corretamente. Verdadeiros negativos são cenas sem objeto corretamente não localizado. Quando o contexto foi utilizado houve uma melhora na rotulação do objeto. Ao utilizar os descritores como probabilidade a priori de cada objeto houve uma elevação nos resultados da localização do objeto alvo. Assim como a melhora no reconhecimento dos objetos dos contextos, ocorreu consequentemente uma melhora no reconhecimento do ambiente passando de 63% a 78% para a escolha padrão copa, e de 79% para 87% para a escolha padrão laboratório. A probabilidade de ser um objeto aumenta proporcionalmente ao número de descritores deste objeto encontrados na imagem. Ou seja, quanto mais descritores, maior a probabilidade daquele objeto estar em cena.

**Tabela 1. Tabela com o resultado da acurácia do experimento**

		Surf	Surf e contexto evidência	Surf e contexto e <i>CRA<sub>CC</sub></i>
<i>ambiente</i>	default Cozinha	—	63%	78%
	default Laboratório	—	79%	87%
<i>objeto</i>	default Copo	17%	31%	38%
	default Porta lápis	29%	40%	41%

## 5. Conclusão

Neste trabalho investigamos reconhecer objetos utilizando informações sobre contexto. Notamos que, ao aplicar um sistema de inferência lógica probabilística, obtivemos uma melhora no reconhecimento de objetos. Este foi um experimento realizado com dados

reais, no qual o ambiente não foi preparado (ou seja sem iluminação especial ou objetos incomuns ou com contraste de objeto e fundo). Os objetos são aqueles que aparecem nas cenas e são típicos de cada ambiente. Podemos notar pelo resultados que continua sendo um desafio trabalhar em ambiente não restritos. Dessa forma, o que estamos testando é um sistema que possa ser utilizado em qualquer ambiente a partir de algum conhecimento a priori dado a um robô doméstico. Utilizamos contexto como a co-ocorrência dos objetos. Contudo podemos notar que, por exemplo, quando o objeto é um porta-lápis ele tem uma relação espacial com os objetos que estão dentro dele. Explorar as relações entre os objetos é um outro modo de definir contexto. Estudar um método de inclusão de um raciocínio espacial do contexto para interpretar cenas será escopo para trabalhos futuros.

## Agradecimentos

Este trabalho recebeu significativo suporte da FAPESP através do Projeto Temático 2008/03995-5 (LogProb). O terceiro autor agradece ao apoio do CNPq (processo 305395/2010-6). A primeira autora é aluna bolsista CAPES e recebeu suporte financeiro da EPUSP.

## Referências

- Baader, F., Calvanese, D., McGuinness, D., Nardi, D., and Patel-Schneider, P. (2002). *Description Logic Handbook*. Cambridge University Press.
- Bar, M. (2004). Visual objects in context. *Nat Rev Neurosci*, 5(8):617–29.
- Bar, M., Ullman, S., et al. (1996). Spatial context in recognition. *PERCEPTION-LONDON*, 25:343–352.
- Bauer, J., Sunderhauf, N., and Protzel, P. (2007). Comparing several implementations of two recently published feature detectors. pages 224–227.
- Bay, H., Ess, A., Tuytelaars, T., and Van Gool, L. (2008). Speeded-up robust features (SURF). *Computer Vision and Image Understanding*, 110(3):346–359.
- Bay, H., Tuytelaars, T., and Van Gool, L. (2006). Surf: Speeded up robust features. *Computer Vision—ECCV 2006*, pages 404–417.
- Bohg, J., Barck-holst, C., Huebner, K., Ralph, M., Rasolzadeh, B., Song, D., and Kragic, D. (2009). Towards grasp-oriented visual perception for humanoid robots. *International Journal of Humanoid Robotics*, 6(3):387–434.
- Bradski, G. (2000). The opencv library. *Dr. Dobb’s Journal of Software Tools*.
- Cozman, F. G. and Polastro, R. B. (2008). Loopy propagation in a probabilistic description logic. In *SUM ’08: Proceedings of the 2nd international conference on Scalable Uncertainty Management*, pages 120–133, Berlin, Heidelberg. Springer-Verlag.
- Fenelon, V., Hummel, B., Santos, P. E., and Cozman, F. G. (2010). Encoding spatial domains with relational bayesian networks. In *UAI 2010*.
- Jain, A., Gupta, A., and Davis, L. S. (2010). Learning what and how of contextual models for scene labeling. In *Proceedings of the 11th European conference on Computer vision: Part IV, ECCV’10*, pages 199–212, Berlin, Heidelberg. Springer-Verlag.

- Korb, K. and Nicholson, A. E. (2003). *Bayesian Artificial Intelligence*. CRC Press, Inc., Boca Raton, FL, USA.
- Lee, B. Y. and , Y. S. (1999). Application of the discrete wavelet transform to the monitoring of tool failure in end milling using the spindle motor current. *The International Journal of Advanced Manufacturing Technology*, 15:238–243. 10.1007/s001700050062.
- Lindeberg, T. (1998). Feature detection with automatic scale selection. *International Journal of Computer Vision*, 30:79–116. 10.1023/A:1008045108935.
- Lowe, D. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110.
- Okabe, T., Kondo, Y., Kitani, K. M., and Sato, Y. (2010). Recognizing multiple object based on co-occurrence of categories. *Progress in informatica*, 7:43–52.
- Parizi, S. N., Laptev, I., and Targhi, A. T. (2009). Modeling image context using object centered grid. In *Proceedings of the 2009 Digital Image Computing: Techniques and Applications*, DICTA '09, pages 476–483, Washington, DC, USA. IEEE Computer Society.
- Polastro, R., Corrêa, F., Cozman, F., and Jr, J. O. (2010). Semantic mapping with a probabilistic description logic. *SBIA*.
- Polastro, R. B. and Cozman, F. G. (2008). Inference in probabilistic ontologies with attributive concept descriptions and nominals. In *URSW'08*, pages –1–1.
- Rabinovich, A.; Vedaldi, A. G. C. W. E. B. S. (2007). Objects in context. *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*.
- Ramisa, A., Aldavert, D., Vasudevan, S., Toledo, R., and de Mántaras, R. L. (2011). Evaluation of three vision based object perception methods for a mobile robot. *CoRR*, abs/1102.0454.
- Ribes, A., Ramisa, A., Lopez de Mantaras, R., and Toledo, R. (2008). Object-based place recognition for mobile robots using panoramas. In *Proceedings of the 2008 conference on Artificial Intelligence Research and Development: Proceedings of the 11th International Conference of the Catalan Association for Artificial Intelligence*, pages 388–397, Amsterdam, The Netherlands, The Netherlands. IOS Press.
- Santos, P., Hummel, B., Fenelon, V., and Cozman, F. G. (2010). Probabilistic logic encoding of spatial domains. In *UniDL 2010*.
- Schmidt-Schauss, M. and Smolka, G. (1991). Attributive concept descriptions with complements. *Artificial Intelligence*, 48(1):1–26.
- Semenovich, D. and Sowmya, A. (2009). A spectral method for context based disambiguation of image annotations. In *Image Processing (ICIP), 2009 16th IEEE International Conference on*, pages 789 –792.
- Shapiro, L. G. and Stockman, G. C. (2001). *Computer Vision*. Tom Robbins.