



Automatic Image Annotation Using Maximum Entropy Model

International Conference on Natural Language Processing

IJCNLP 2005: Natural Language Processing – IJCNLP 2005 pp 34-45 | Cite as

Wei Li (1)

Maosong Sun (1)

1. State Key Lab of Intelligent Technology and Systems, Department of Computer Science and Technology, Tsinghua University, , Beijing, China

Conference paper

[1 Citations](#)

[937 Downloads](#)

Part of the [Lecture Notes in Computer Science](#) book series (LNCS, volume 3651)

Abstract

Automatic image annotation is a newly developed and promising technique to provide semantic image retrieval via text descriptions. It concerns a process of automatically labeling the image contents with a pre-defined set of keywords which are exploited to represent the image semantics. A Maximum Entropy Model-based approach to the task of automatic image annotation is proposed in this paper. In the phase of training, a basic visual vocabulary consisting of blob-tokens to describe the image content is generated at first; then the statistical relationship is modeled between the blob-tokens and keywords by a Maximum Entropy Model constructed from the training set of labeled images. In the phase of annotation, for an unlabeled image, the most likely associated keywords are predicted in terms of the blob-token set extracted from the given image. We carried out experiments on a medium-sized image collection with about 5000 images from Corel Photo CDs. The experimental results demonstrated that the annotation performance of this method outperforms some traditional annotation methods by about 8% in mean precision, showing a potential of the Maximum Entropy Model in the task of automatic image annotation.

Keywords

We use cookies to personalise content and ads, to provide social media features and to analyse our traffic. We also share information about your use of our site with our social media, advertising and analytics partners in accordance with our [Privacy Statement](#). You can manage your preferences in [Manage Cookies](#).

Unable to display preview. [Download preview PDF.](#)

References

1. Barnard, K., Dyugulu, P., de Freitas, N., Forsyth, D., Blei, D., Jordan, M.I.: Matching words and pictures. *Journal of Machine Learning Research* 3, 1107–1135 (2003)
[zbMATH](http://www.emis.de/MATH-item?1061.68174) (<http://www.emis.de/MATH-item?1061.68174>)
[CrossRef](https://doi.org/10.1162/153244303322533214) (<https://doi.org/10.1162/153244303322533214>)
[Google Scholar](http://scholar.google.com/scholar_lookup?title=Matching%20words%20and%20pictures&author=K..%20Barnard&author=P..%20Dyugulu&author=N..%20Freitas&author=D..%20Forsyth&author=D..%20Blei&author=M.I..%20Jordan&journal=Journal%20of%20Machine%20Learning%20Research&volume=3&pages=1107-1135&publication_year=2003) (http://scholar.google.com/scholar_lookup?title=Matching%20words%20and%20pictures&author=K..%20Barnard&author=P..%20Dyugulu&author=N..%20Freitas&author=D..%20Forsyth&author=D..%20Blei&author=M.I..%20Jordan&journal=Journal%20of%20Machine%20Learning%20Research&volume=3&pages=1107-1135&publication_year=2003)
2. Duygulu, P., Barnard, K., de Freitas, N., Forsyth, D.: Object recognition as machine translation: Learning a lexicon from a fixed image vocabulary. In: *Seventh European Conf. on Computer Vision*, pp. 97–112 (2002)
[Google Scholar](https://scholar.google.com/scholar?q=Duygulu%2C%20P..%2C%20Barnard%2C%20K..%2C%20de%20Freitas%2C%20N..%2C%20Forsyth%2C%20D..%3A%20Object%20recognition%20as%20machine%20translation%3A%20Learning%20a%20lexicon%20from%20a%20fixed%20image%20vocabulary.%20In%3A%20Seventh%20European%20Conf.%20on%20Computer%20Vision%2C%20pp.%2097%E2%80%93112%20%282002%29) (<https://scholar.google.com/scholar?q=Duygulu%2C%20P..%2C%20Barnard%2C%20K..%2C%20de%20Freitas%2C%20N..%2C%20Forsyth%2C%20D..%3A%20Object%20recognition%20as%20machine%20translation%3A%20Learning%20a%20lexicon%20from%20a%20fixed%20image%20vocabulary.%20In%3A%20Seventh%20European%20Conf.%20on%20Computer%20Vision%2C%20pp.%2097%E2%80%93112%20%282002%29>)
3. Jeon, J., Lavrenko, V., Manmatha, R.: Automatic image annotation and retrieval using cross-media relevance models. In: *Proceedings of the 26th Intl. SIGIR Conf*, pp. 119–126 (2003)
[Google Scholar](https://scholar.google.com/scholar?q=Jeon%2C%20J..%2C%20Lavrenko%2C%20V..%2C%20Manmatha%2C%20R..%3A%20Automatic%20image%20annotation%20and%20retrieval%20using%20cross-media%20relevance%20models.%20In%3A%20Proceedings%20of%20the%20) (<https://scholar.google.com/scholar?q=Jeon%2C%20J..%2C%20Lavrenko%2C%20V..%2C%20Manmatha%2C%20R..%3A%20Automatic%20image%20annotation%20and%20retrieval%20using%20cross-media%20relevance%20models.%20In%3A%20Proceedings%20of%20the%20>

[Manage Cookies](#)

✓ OK

We use cookies to personalise content and ads, to provide social media features and to analyse our traffic. We also share information about your use of our site with our social media, advertising and analytics partners in accordance with our [Privacy Statement](#). You can manage your preferences in [Manage Cookies](#).

%20Workshop%20on%20Multimedia%20Intelligent%20Storage%20and%20Retrieval%20Management%20%281999%29)

5. Chang, E., Goh, K., Sychay, G., Wu, G.: CBSA: Content-based soft annotation for multimodal image retrieval using bayes point machines. *IEEE Transactions on Circuits and Systems for Video Technology Special Issue on Conceptual and Dynamical Aspects of Multimedia Content Descriptions* 13(1), 26–38 (2003) [Google Scholar](http://scholar.google.com/scholar_lookup?title=CBSA%3A%20Content-based%20soft%20annotation%20for%20multimodal%20image%20retrieval%20using%20bayes%20point%20machines&author=E.%20Chang&author=K.%20Goh&author=G.%20Sychay&author=G.%20Wu&journal=IEEE%20Transactions%20on%20Circuits%20and%20Systems%20for%20Video%20Technology%20Special%20Issue%20on%20Conceptual%20and%20Dynamical%20Aspects%20of%20Multimedia%20Content%20Descriptions&volume=13&issue=1&pages=26-38&publication_year=2003) (http://scholar.google.com/scholar_lookup?title=CBSA%3A%20Content-based%20soft%20annotation%20for%20multimodal%20image%20retrieval%20using%20bayes%20point%20machines&author=E.%20Chang&author=K.%20Goh&author=G.%20Sychay&author=G.%20Wu&journal=IEEE%20Transactions%20on%20Circuits%20and%20Systems%20for%20Video%20Technology%20Special%20Issue%20on%20Conceptual%20and%20Dynamical%20Aspects%20of%20Multimedia%20Content%20Descriptions&volume=13&issue=1&pages=26-38&publication_year=2003)
6. Shi, J., Malik, J.: Normalized cuts and image segmentation. *IEEE Transactions On Pattern Analysis and Machine Intelligence* 22(8), 888–905 (2000) [CrossRef](https://doi.org/10.1109/34.868688) (<https://doi.org/10.1109/34.868688>) [Google Scholar](http://scholar.google.com/scholar_lookup?title=Normalized%20cuts%20and%20image%20segmentation&author=J.%20Shi&author=J.%20Malik&journal=IEEE%20Transactions%20On%20Pattern%20Analysis%20and%20Machine%20Intelligence&volume=22&issue=8&pages=888-905&publication_year=2000) (http://scholar.google.com/scholar_lookup?title=Normalized%20cuts%20and%20image%20segmentation&author=J.%20Shi&author=J.%20Malik&journal=IEEE%20Transactions%20On%20Pattern%20Analysis%20and%20Machine%20Intelligence&volume=22&issue=8&pages=888-905&publication_year=2000)
7. Berger, A., Pietra, S., Pietra, V.: A maximum entropy approach to natural language processing. In: *Computational Linguistics*, pp. 39–71 (1996) [Google Scholar](https://scholar.google.com/scholar?q=Berger%2C%A.%2C%20Pietra%2C%20S.%2C%20Pietra%2C%20V.%3A%20A%20maximum%20entropy%20approach%20to%20natural%20language%20processing.%20In%3A%20Computational%20Linguistics%2C%20pp.%2039%20-%2071%20%281996%29) (<https://scholar.google.com/scholar?q=Berger%2C%A.%2C%20Pietra%2C%20S.%2C%20Pietra%2C%20V.%3A%20A%20maximum%20entropy%20approach%20to%20natural%20language%20processing.%20In%3A%20Computational%20Linguistics%2C%20pp.%2039%20-%2071%20%281996%29>)
8. Li, J., Wang, J.A.: Automatic linguistic indexing of pictures by a statistical modeling approach. *IEEE Transactions on PAMI* 25(10), 1075–1088 (2003) [Google Scholar](http://scholar.google.com/scholar_lookup?title=Automatic%20linguistic%20indexing%20of%20pictures%20by%20a%20statistical%20modeling%20approach&author=J.%20Li&author=J.A.%20Wang&journal=IEEE%20Transactions%20on%20PAMI&volume=25&issue=10&pages=1075-1088&publication_year=2003) (http://scholar.google.com/scholar_lookup?title=Automatic%20linguistic%20indexing%20of%20pictures%20by%20a%20statistical%20modeling%20approach&author=J.%20Li&author=J.A.%20Wang&journal=IEEE%20Transactions%20on%20PAMI&volume=25&issue=10&pages=1075-1088&publication_year=2003)
9. Jeon, J., Manmatha, R.: Using maximum entropy for automatic image annotation. In: *Proceedings of third international conference on image and video retrieval*, pp. 24–31 (2004)

We use cookies to personalise content and ads, to provide social media features and to analyse our traffic. We also share information about your use of our site with our social media, advertising and analytics partners in accordance with our [Privacy Statement](#). You can manage your preferences in [Manage Cookies](#).

maximum entropy parameter estimation. In: Proceedings of the 2006th Workshop on Computational Language Learning (2006)

11. Lavrenko, V., Manmatha, R., Jeon, J.: A model for learning the semantics of pictures. In: Proceedings of the 16th Annual Conference on Neural Information Processing Systems (2004)
[Google Scholar](https://scholar.google.com/scholar?q=Lavrenko%2C%20V.%2C%20Manmatha%2C%20R.%2C%20Jeon%2C%20J.%3A%20A%20model%20for%20learning%20the%20semantics%20of%20pictures.%20In%3A%20Proceedings%20of%20the%2016th%20Annual%20Conference%20on%20Neural%20Information%20Processing%20Systems%20%282004%29) (https://scholar.google.com/scholar?q=Lavrenko%2C%20V.%2C%20Manmatha%2C%20R.%2C%20Jeon%2C%20J.%3A%20A%20model%20for%20learning%20the%20semantics%20of%20pictures.%20In%3A%20Proceedings%20of%20the%2016th%20Annual%20Conference%20on%20Neural%20Information%20Processing%20Systems%20%282004%29)

Copyright information

© Springer-Verlag Berlin Heidelberg 2005

About this paper

Cite this paper as:

Li W., Sun M. (2005) Automatic Image Annotation Using Maximum Entropy Model. In: Dale R., Wong KF., Su J., Kwong O.Y. (eds) Natural Language Processing – IJCNLP 2005. IJCNLP 2005. Lecture Notes in Computer Science, vol 3651. Springer, Berlin, Heidelberg

DOI https://doi.org/10.1007/11562214_4

Publisher Name Springer, Berlin, Heidelberg

Print ISBN 978-3-540-29172-5

Online ISBN 978-3-540-31724-1

eBook Packages [Computer Science](#)

[Buy this book on publisher's site](#)

[Reprints and Permissions](#)

Personalised recommendations

SPRINGER NATURE

© 2018 Springer Nature Switzerland AG. Part of [Springer Nature](#).

Not logged in Not affiliated 110.182.192.174

We use cookies to personalise content and ads, to provide social media features and to analyse our traffic. We also share information about your use of our site with our social media, advertising and analytics partners in accordance with our [Privacy Statement](#). You can manage your preferences in [Manage Cookies](#).