WVL '13

NAACL HLT 2013
Workshop on Vision and Language

Proceedings of the Workshop

14 June 2013
Westin Peachtree Plaza
Atlanta, Georgia, USA

209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-937284-47-3

Introduction

Welcome to the HLT NAACL Workshop on Vision and Language (WVL'13).

There is an increasing amount of research at the interfaces of speech and language processing and computer vision, computer graphics, robotics and information retrieval which aims to develop systems that automatically generate descriptions of images or videos, or generate images based on natural language descriptions, acquire and understand language in a perceptually grounded, visual context, or perform language-based image search.

Since the main purpose of this workshop is to bring researchers from these communities together, the workshop will mostly consist of invited talks, both by NLP and computer vision students who are working in the area, as well as by established researchers from academia and industry.

Organizers:

Julia Hockenmaier, University of Illinois at Urbana-Champaign
Tamara Berg, Stony Brook University

Program Committee:

Samy Bengio, Google
Alexander C Berg, Stony Brook University
Yejin Choi, Stony Brook University
Bill Dolan, Microsoft Research, Redmond
Jacob Eisenstein, Georgia Institute of Technology
Desmond Elliott, University of Edinburgh
Michel Galley, Microsoft Research, Redmond
Kristen Grauman, University of Texas, Austin
John Kelleher, Dublin Institute of Technology
Mirella Lapata, University of Edinburgh
Margaret Mitchell, Johns Hopkins University
Ray Mooney, University of Texas, Austin
Owen Rambow, Columbia University
Richard Sproat, Google

Table of Contents

Workshop Program

Friday, June 14, 2013

Session 1

8:45–9:00 Opening Remarks

9:00–10:00 Tutorial: *Computational Visual Recognition for NLP*
Alexander C Berg (Stony Brook University)

10:00–10:30 Invited talk: *Modality Selection for Multimedia Summarization*
Florian Metze (Carnegie Mellon University)

10:30–11:00 Coffee break

Session 2

11:00–11:20 *Annotation of Online Shopping Images without Labeled Training Examples*
Rebecca Mason and Eugene Charniak

11:20–11:40 *Generating Natural-Language Video Descriptions Using Text-Mined Knowledge*
Niveda Krishnamoorthy, Girish Malkarnenkar, Raymond Mooney,
Kate Saenko and Sergio Guadarrama

11:40–12:00 *Learning Hierarchical Linguistic Descriptions of Visual Datasets*
Roni Mittelman, Min Sun, Benjamin Kuipers and Silvio Savarese

12:00–12:30 Invited talk: *Joint Learning of Word Meanings and Image Tasks*
Jason Weston (Google)

12:30–2:00 Lunch Break

Friday, June 14, 2013 (continued)

Session 3

2:00–2:15 Invited student talk:
Communicating with an Image Retrieval System via Relative Attributes
Adriana Kovashka and Kristen Grauman (University of Texas at Austin)

2:15–2:30 Invited student talk: *Identifying Visual Attributes for Object Recognition*
Caglar Tirkaz, Jacob Eisenstein, Berrin Yanikoglu and Metin Sezgin
(Sabanci University, Georgia Institute of Technology, Koc University)

2:30–2:45 Invited student talk: *Generating Visual Descriptions from Feature Norms of Actions, Attributes, Classes and Parts*
Mark Yatskar and Luke Zettlemoyer (University of Washington)

2:45–3:00 Invited student talk: *Bayesian modeling of scenes and captions*
Luca del Pero and Kobus Barnard (University of Arizona)

3:00–3:15 Invited student talk: *Data-Driven Generation of Image Descriptions*
Vicente Ordonez and Tamara Berg (Stony Brook University)

3:15–3:30 Invited student talk: *Framing image description as a retrieval problem*
Micah Hodosh, Peter Young and Julia Hockenmaier
(University of Illinois at Urbana-Champaign)

Session 4

4:00–4:30 Invited talk: *Multimodal Semantics at CLIC*
Elia Bruni (University of Trento)

4:30–5:00 Invited talk: *Generating and Generalizing Image Captions*
Yejin Choi (Stony Brook University)

5:00–5:30 Invited talk: *Generating Descriptions of Visible Objects*
Margaret Mitchell (Johns Hopkins University)

5:30–6:00 Panel discussion
Julia Hockenmaier and Tamara Berg