

**3rd Annual Meeting of the EPSRC
Network on Vision and Language and
the 1st Technical Meeting of the
European Network on Integrating
Vision and Language**

(V&L Net 2014)

Held at COLING 2014

**Dublin, Ireland
23 August 2014**

ISBN: 978-1-63439-222-8

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2014) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2014)

For permission requests, please contact the Association for Computational Linguistics
at the address below.

Association for Computational Linguistics
209 N. Eighth Street
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006
Fax: 1-570-476-0860

acl@aclweb.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2634
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

<i>The Effect of Sensor Errors in Situated Human-Computer Dialogue</i> Niels Schütte, John Kelleher and Brian Mac Namee	1
<i>Joint Navigation in Commander/Robot Teams: Dialog & Task Performance When Vision is Bandwidth-Limited</i> Douglas Summers-Stay, Taylor Cassidy and Clare Voss	9
<i>TUHOI: Trento Universal Human Object Interaction Dataset</i> Dieu-Thu Le, Jasper Uijlings and Raffaella Bernardi	17
<i>Concept-oriented labelling of patent images based on Random Forests and proximity-driven generation of synthetic data</i> Dimitris Liparas, Anastasia Moutzidou, Stefanos Vrochidis and Ioannis Kompatsiaris	25
<i>Exploration of functional semantics of prepositions from corpora of descriptions of visual scenes</i> Simon Dobnik and John Kelleher	33
<i>A Poodle or a Dog? Evaluating Automatic Image Annotation Using Human Descriptions at Different Levels of Granularity</i> Josiah Wang, Fei Yan, Ahmet Aker and Robert Gaizauskas	38
<i>Key Event Detection in Video using ASR and Visual Data</i> Niraj Shrestha, Aparna N. Venkitasubramanian and Marie-Francine Moens	46
<i>Twitter User Gender Inference Using Combined Analysis of Text and Image Processing</i> Shigeyuki Sakaki, Yasuhide Miura, Xiaojun Ma, Keigo Hattori and Tomoko Ohkuma	54
<i>Semantic and geometric enrichment of 3D geo-spatial models with captioned photos and labelled illustrations</i> Chris Jones, Paul Rosin and Jonathan Slade	62
<i>Weakly supervised construction of a repository of iconic images</i> Lydia Weiland, Wolfgang Effelsberg and Simone Paolo Ponzetto	68
<i>Cross-media Cross-genre Information Ranking based on Multi-media Information Networks</i> Tongtao Zhang, Haibo Li, Hongzhao Huang, Heng Ji, Min-Hsuan Tsai, Shen-Fu Tsai and Thomas Huang	74
<i>Speech-accompanying gestures in Russian: functions and verbal context</i> Yulia Nikolaeva	82
<i>DALES: Automated Tool for Detection, Annotation, Labelling, and Segmentation of Multiple Objects in Multi-Camera Video Streams</i> Mohammad Bhat and Joanna Isabelle Olszewska	87
<i>A Hybrid Segmentation of Web Pages for Vibro-Tactile Access on Touch-Screen Devices</i> Waseem SAFI, Fabrice Maurel, Jean-Marc Routoure, Pierre Beust and Gaël Dias	95
<i>Expression Recognition by Using Facial and Vocal Expressions</i> Gholamreza Anbarjafari and Alvo Aabloo	103

<i>Formulating Queries for Collecting Training Examples in Visual Concept Classification</i> Kevin McGuinness, Feiyang Hu, Rami Albatat and Alan Smeaton	106
<i>Towards Succinct and Relevant Image Descriptions</i> Desmond Elliott	109
<i>Coloring Objects: Adjective-Noun Visual Semantic Compositionality</i> Dat Tien Nguyen, Angeliki Lazaridou and Raffaella Bernardi	112
<i>Multi-layered Image Representation for Image Interpretation</i> Marina Ivacic-Kos, Miran Pobar and Ivo Ipsic	115
<i>The Last 10 Metres: Using Visual Analysis and Verbal Communication in Guiding Visually Impaired Smartphone Users to Entrances</i> Anja Belz and Anil Bharath	118
<i>Keyphrase Extraction using Textual and Visual Features</i> Yaakov HaCohen-Kerner, Stefanos Vrochidis, Dimitris Liparas, Anastasia Moutzidou and Ioannis Kompatsiaris	121
<i>Towards automatic annotation of communicative gesturing</i> Kristiina Jokinen and Graham Wilcock	124