



Universiteit  
Leiden  
The Netherlands

## Large scale visual search

Wu, S.

### Citation

Wu, S. (2016, December 22). *Large scale visual search*. Retrieved from <https://hdl.handle.net/1887/45135>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/45135>

**Note:** To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/45135> holds various files of this Leiden University dissertation.

**Author:** Wu, S.

**Title:** Large scale visual search

**Issue Date:** 2016-12-22

# Large Scale Visual Search

## Proefschrift

ter verkrijging van  
de graad van Doctor aan de Universiteit Leiden,  
op gezag van Rector Magnificus prof.mr. C.J.J.M. Stolker,  
volgens besluit van het College voor Promoties  
te verdedigen op donderdag 22 december 2016  
klokke 16.15 uur

door

**Song Wu**

geboren te Sichuan, China  
in 1985

## Promotiecommissie

Promotor: Prof. Dr. J.N. Kok  
Co-promotor: Dr. M.S. Lew  
Overige leden: Prof. Dr. A. Plaat  
Prof. Dr. W. Kraaij  
Prof. Dr. T.H.W. Bäck  
Prof. Dr. C. Griwodz (University of Oslo)  
Prof. Dr. M. Larson (Delft University of Technology)

Copyright © 2016 Song Wu All Rights Reserved

ISBN/AEN 9789463321174

Cover photo: The cover photo shows the flowchart of the proposed deep binary codes used for large scale visual search.

This research is financially supported by the China Scholarship Council (CSC), Grant No. 201206990003.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Salient Point Methods . . . . .	2
1.2	Visual Word based Image Search . . . . .	6
1.3	Convolutional Neural Networks . . . . .	8
1.4	Thesis Overview . . . . .	12
<b>2</b>	<b>A Comprehensive Evaluation of Salient Point Methods</b>	<b>17</b>
2.1	Introduction . . . . .	18
2.2	Background . . . . .	19
2.3	Overview of Evaluated Salient Point Methods . . . . .	22
2.3.1	SIFT (detector/descriptor) . . . . .	23
2.3.2	SURF (detector/descriptor) . . . . .	25
2.3.3	MSER (detector) . . . . .	25
2.3.4	HESSIAN-AFFINE (detector) . . . . .	26
2.3.5	FAST (detector) . . . . .	26
2.3.6	CenSurE (detector) . . . . .	27
2.3.7	GFTT (detector) . . . . .	27
2.3.8	KAZE (detector) . . . . .	28
2.3.9	BRIEF (descriptor) . . . . .	28
2.3.10	ORB (detector/descriptor) . . . . .	29
2.3.11	BRISK (detector/descriptor) . . . . .	30
2.3.12	FREAK (descriptor) . . . . .	30
2.3.13	BinBoost (descriptor) . . . . .	31
2.3.14	LATCH (descriptor) . . . . .	31

## CONTENTS

---

2.4	Fully Affine Space Framework . . . . .	31
2.5	Experimental Setup . . . . .	34
2.5.1	Datasets . . . . .	34
2.5.2	Evaluation Criteria . . . . .	34
2.6	Results and Discussions . . . . .	35
2.6.1	Detector Evaluation . . . . .	35
2.6.2	Descriptor Evaluation . . . . .	38
2.6.3	Affine Invariant Evaluation . . . . .	41
2.6.3.1	Parameter of $K$ in $K$ -order $NNDR$ . . . . .	44
2.6.3.2	Correspondence Matching Using the Framework of Fully Affine Space . . . . .	45
2.6.3.3	Computational Cost and Memory Requirement . . . . .	49
2.7	Conclusions . . . . .	50
<b>3</b>	<b>RIFF: Retina-inspired Invariant Fast Feature Descriptor</b>	<b>51</b>
3.1	Introduction . . . . .	52
3.2	Discriminate RIFF Local Descriptor . . . . .	54
3.2.1	Retina Sampling Pattern Review . . . . .	54
3.2.2	Descriptor Generation . . . . .	55
3.2.2.1	Orientation Estimation . . . . .	55
3.2.2.2	Descriptor Generation . . . . .	56
3.2.2.3	Discriminative Strategy . . . . .	57
3.3	Visual Word Model based Image Search . . . . .	58
3.4	Experimental Results . . . . .	61
3.4.1	Datasets and Evaluation Criteria . . . . .	62
3.4.2	Evaluation of Image Copy Detection . . . . .	63
3.4.2.1	Evaluation of Time and Storage . . . . .	64
3.4.2.2	Evaluation of Search Accuracy . . . . .	64
3.5	Conclusions . . . . .	67
<b>4</b>	<b>Deep Binary Codes for Large Scale Image Retrieval</b>	<b>69</b>
4.1	Introduction . . . . .	70
4.2	Related Work . . . . .	74
4.3	Proposed Approach . . . . .	75

4.3.1	Generating Deep Binary Codes . . . . .	75
4.3.2	Spatial Cross-Summing . . . . .	78
4.3.3	Dynamic Late Fusion . . . . .	79
4.4	Experiments and Setup . . . . .	81
4.4.1	Datasets . . . . .	82
4.4.2	Evaluation of Deep Convolutional Feature Representation .	83
4.4.3	Performance of Deep Binary Codes . . . . .	84
4.4.4	Comparison with Hashing Learning Approaches . . . . .	85
4.4.5	Evaluation of the Late Fusion Scheme . . . . .	87
4.4.6	Performance on Large Scale Image Search . . . . .	90
4.4.7	Comparison with state-of-the-art . . . . .	92
4.5	Conclusions . . . . .	92
<b>5</b>	<b>Comparison of Information Loss Architectures in CNNs</b>	<b>93</b>
5.1	Introduction . . . . .	94
5.2	Related Work . . . . .	96
5.3	Convolutional Neural Networks Classification . . . . .	97
5.4	Integration Architecture Network . . . . .	98
5.4.1	Concatenate Architecture Network . . . . .	98
5.4.2	Weighted Integration Architecture Network . . . . .	99
5.5	Experimental Results . . . . .	102
5.5.1	Datasets . . . . .	102
5.5.2	Details of Weighted Integration Architecture . . . . .	103
5.5.3	Evaluation Results . . . . .	104
5.6	Conclusions . . . . .	106
<b>6</b>	<b>Conclusions</b>	<b>107</b>
6.1	Conclusions . . . . .	107
6.2	Future Work . . . . .	109
	<b>Bibliography</b>	<b>113</b>
	<b>English Summary</b>	<b>133</b>
	<b>Nederlandse Samenvatting</b>	<b>135</b>

## CONTENTS

---

Acknowledgements	137
Curriculum Vitae	139