

## MASTER'S THESIS

### Learning hash codes for multimedia retrieval

Chen, Junjie

*Date of Award:*  
2019

[Link to publication](#)

#### **General rights**

Copyright and intellectual property rights for the publications made accessible in HKBU Scholars are retained by the authors and/or other copyright owners. In addition to the restrictions prescribed by the Copyright Ordinance of Hong Kong, all users and readers must also observe the following terms of use:

- Users may download and print one copy of any publication from HKBU Scholars for the purpose of private study or research
- Users cannot further distribute the material or use it for any profit-making activity or commercial gain
- To share publications in HKBU Scholars with others, users are welcome to freely distribute the permanent URL assigned to the publication

# Abstract

The explosive growth of multimedia data in online media repositories and social networks has led to the high demand of fast and accurate services for large-scale multimedia retrieval. Hashing, due to its effectiveness in coding high-dimensional data into a low-dimensional binary space, has been considered to be effective for the retrieval application. Despite the progress that has been made recently, how to learn the optimal hashing models which can make the best trade-off between the retrieval efficiency and accuracy remains to be open research issues. This thesis research aims to develop hashing models which are effective for image and video retrieval. An unsupervised hashing model called APHash is first proposed to learn hash codes for images by exploiting the distribution of data. To reduce the underlying computational complexity, a methodology that makes use of an asymmetric similarity matrix is explored and found effective. In addition, the deep learning approach to learn hash codes for images is also studied. In particular, a novel deep model called DeepQuan which tries to incorporate product quantization methods into an unsupervised deep model for the learning. Other than adopting only the quadratic loss as the optimization objective like most of the related deep models, DeepQuan optimizes the data representations and their quantization codebooks to explore the clustering structure of the underlying data manifold where the introduction of a weighted triplet loss into the learning objective is found to be effective. Furthermore, the case with some labeled data available for the learning is also considered. To alleviate the high training cost (which is especially crucial given a large-scale database), another hashing model named Similarity Preserving Deep Asymmetric

Quantization (SPDAQ) is proposed for both image and video retrieval where the compact binary codes and quantization codebooks for all the items in the database can be explicitly learned in an efficient manner. All the aforementioned hashing methods proposed have been rigorously evaluated using benchmark datasets and found to outperform the related state-of-the-art methods.

**Keywords:** hashing, quantization, multimedia retrieval, image retrieval

# Table of Contents

<b>Declaration</b>	<b>i</b>
<b>Abstract</b>	<b>ii</b>
<b>Acknowledgements</b>	<b>iv</b>
<b>Table of Contents</b>	<b>v</b>
<b>List of Tables</b>	<b>viii</b>
<b>List of Figures</b>	<b>x</b>
<b>Chapter 1 Introduction</b>	<b>1</b>
1.1 Problem Definition . . . . .	1
1.2 Thesis Outline . . . . .	4
<b>Chapter 2 Related Work</b>	<b>7</b>
2.1 Nearest Neighbor Search . . . . .	7
2.2 Data-independent Hashing . . . . .	8
2.3 Data-dependent Hashing . . . . .	9
2.3.1 Deep Learning . . . . .	10
2.3.2 Supervised Hashing . . . . .	10
2.3.3 Unsupervised Hashing . . . . .	17
<b>Chapter 3 APHash: Anchor-based Probability Hashing for Image Retrieval</b>	<b>23</b>

3.1	Motivation . . . . .	23
3.2	Methodology . . . . .	26
3.2.1	Problem Formulation . . . . .	26
3.2.2	Optimization . . . . .	29
3.2.3	Hash Function Learning . . . . .	30
3.3	Experimental Results . . . . .	31
3.3.1	Datasets and Experimental Setup . . . . .	31
3.3.2	Comparison with State-of-the-art Methods . . . . .	37
3.3.3	Model Analysis . . . . .	39
3.4	Conclusions . . . . .	41
 <b>Chapter 4 DeepQuan: Learning Deep Unsupervised Binary Codes for Image Retrieval</b>		<b>42</b>
4.1	Motivation . . . . .	42
4.2	Methodology . . . . .	45
4.2.1	Model Formulation . . . . .	45
4.2.2	Approximate Nearest Neighbor Search . . . . .	49
4.2.3	Alternating Optimization . . . . .	50
4.3	Experimental Results . . . . .	52
4.3.1	Datasets and Experimental Setting . . . . .	52
4.3.2	Performance Comparison . . . . .	55
4.3.3	Empirical Analysis of DeepQuan . . . . .	55
4.3.4	Parameters Analysis . . . . .	61
4.4	Conclusions . . . . .	62
 <b>Chapter 5 SPDAQ: Similarity Preserving Deep Asymmetric Quanti- zation for Multimedia Retrieval</b>		<b>63</b>
5.1	Motivation . . . . .	63
5.2	Methodology . . . . .	64
5.2.1	Notation . . . . .	65

5.2.2	Model Formulation . . . . .	67
5.2.3	Optimization . . . . .	72
5.2.4	Out-of-Sample Extension . . . . .	76
5.3	Experimental Results on Image Retrieval . . . . .	77
5.3.1	Experimental Settings . . . . .	77
5.3.2	Retrieval Accuracy . . . . .	84
5.3.3	Efficiency for Model Training . . . . .	84
5.3.4	Variants of SPDAQ . . . . .	85
5.4	Experimental Results on Video Retrieval . . . . .	86
5.4.1	Formulation . . . . .	86
5.4.2	Experimental Settings and Datasets . . . . .	87
5.4.3	Experimental Results . . . . .	93
5.5	Conclusions . . . . .	94
<b>Chapter 6 Conclusions and Future Works</b>		<b>96</b>
<b>Bibliography</b>		<b>98</b>
<b>Curriculum Vitae</b>		<b>111</b>