澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU

ASYMMETRIC SCALABLE CROSS-MODAL HASHING

Published in: 2023 IEEE International Conference on Image Processing (ICIP)

um 澳大

Abstract

Cross-modal hashing is a practical approach to solving the problem of large-scale multimedia retrieval. However, there are still specific issues that the current methods cannot solve, such as how to construct binary codes rather than relax them to continuity effectively and how to prevent nxn problem. This paper proposes a novel Asymmetric Scalable Cross-Modal Hashing (ASCMH) to address these issues. It learns a common latent space from the kernelized features of different modalities. It then transforms the similarity matrix optimization to a distance-distance difference minimization problem with the help of semantic labels and common latent space. Additionally, we use an orthogonal constraint of label information to construct hash codes necessary for search accuracy. Extensive experiments on three benchmark datasets show that our ASCMH outperforms the SOTA cross-modal hashing methods.

澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU

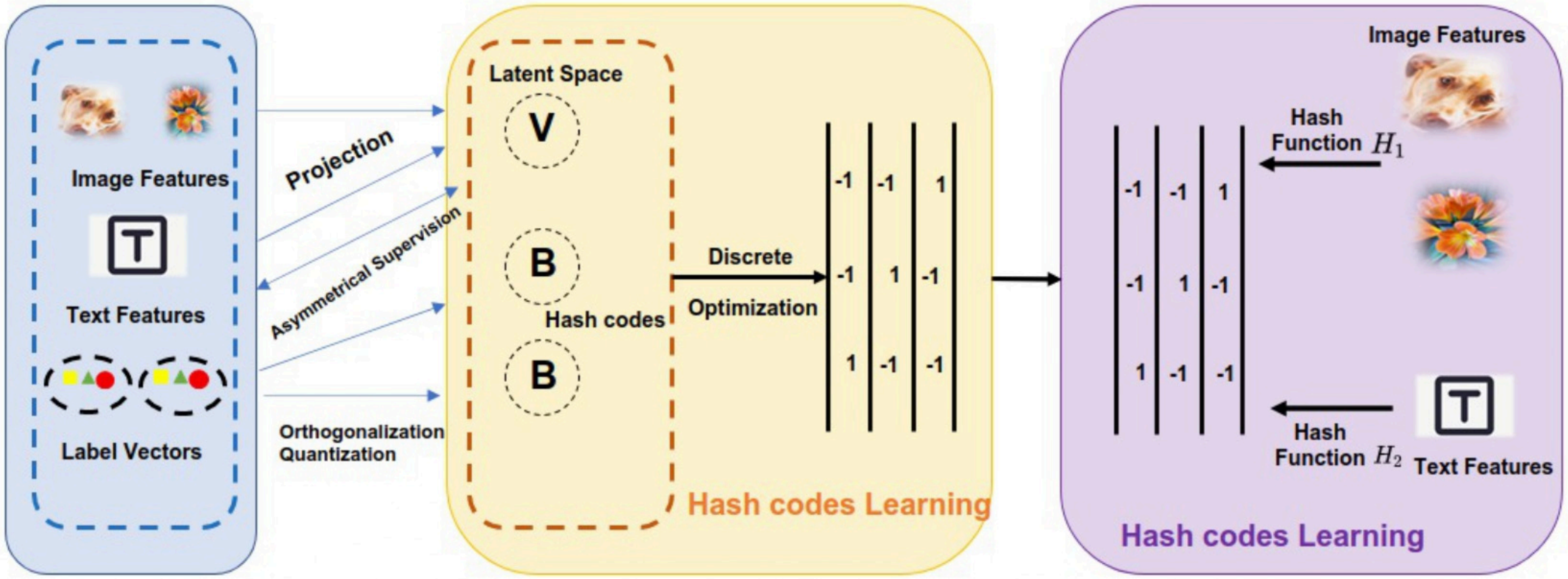
Introduction

In the era of big data, efficient and accurate retrieval of multimedia content has become a crucial task. Cross-modal hashing, which aims to map multimedia data from different modalities into a common binary code space, has gained significant attention due to its ability to facilitate fast and scalable retrieval. In this paper, we present a novel approach called Asymmetric Scalable Cross-Modal Hashing (ASCMH) to address the challenges associated with constructing binary codes effectively and preventing the $n \times n$ problem in large-scale multimedia retrieval.

Problem Statement

The construction of binary codes is a critical aspect of cross-modal hashing. Traditional methods often suffer from information loss and suboptimal performance. Additionally, the $n \times n$ problem arises when dealing with large-scale datasets, where the computational complexity increases exponentially. To overcome these challenges, ASCMH introduces a novel framework that leverages semantic labels to optimize similarity matrices and improve search accuracy.

Methodology



The framework of the proposed Asymmetric Scalable Cross-Modal Hashing (ASCMH).

Methodology

ASCMH learns a common latent space by jointly optimizing the similarity matrices of different modalities. This is achieved by incorporating semantic labels, which provide valuable information about the underlying relationships between multimedia data. The optimization process aims to minimize the semantic gap and maximize the discriminative power of the learned binary codes. By utilizing a scalable hashing strategy, ASCMH effectively addresses the $n \times n$ problem and enables efficient retrieval in large-scale datasets.

Experimental Evaluation

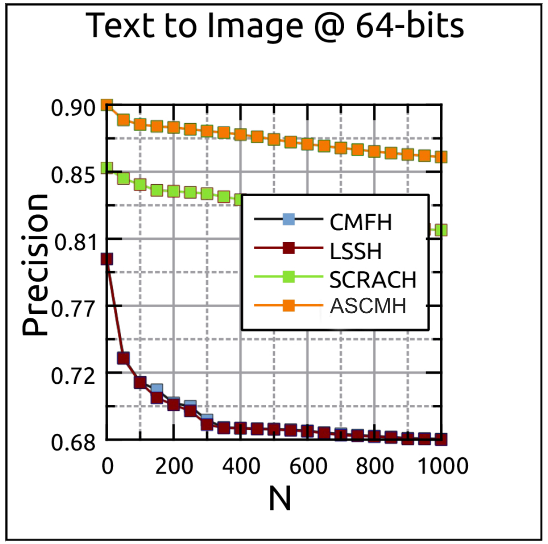
To evaluate the performance of ASCMH, extensive experiments were conducted on three benchmark datasets: Wiki, MIRFlickr25k, and NUS-wide. Mean Average Precision (mAP) and Top-N precision were used as assessment criteria, following the methodology of previous studies. The experiments involved comparing ASCMH with other state-of-the-art cross-modal hashing methods, and the results were averaged over 20 runs to ensure robustness.

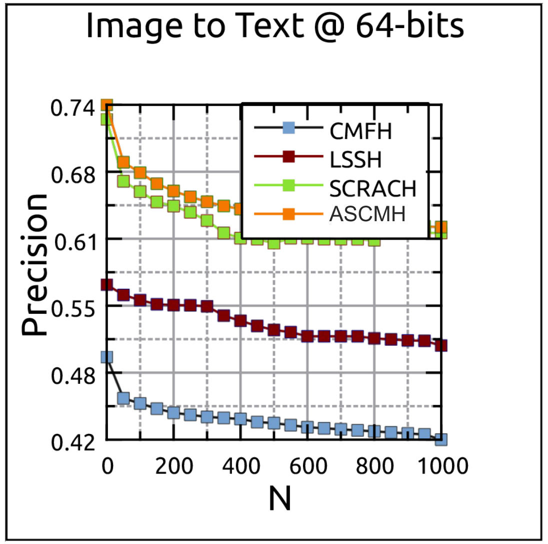
Experimental Result

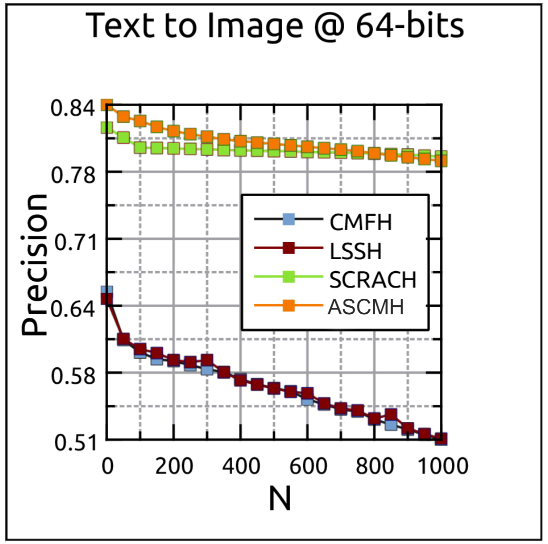
Table 1. The mAP values on three datasets for various code lengths

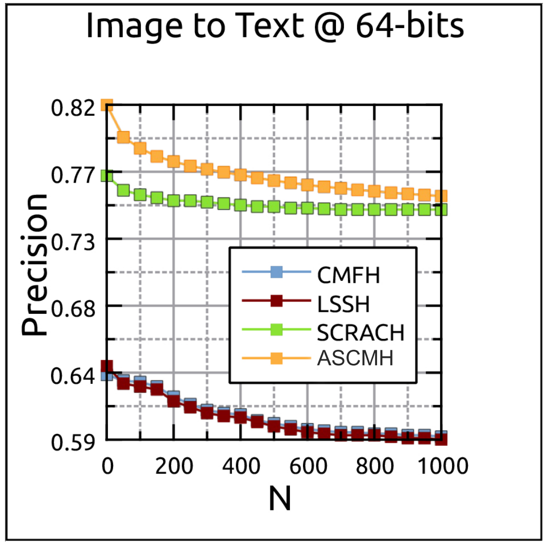
Task	Method	Wiki				MIRFlickr25k				NUS-wide			
		16 bits	32 bits	64 bits	128 bits	16 bits	32 bits	64 bits	128 bits	16 bits	32 bits	32 bits	128 bits
Img to Text	CMFH	0.1664	0.165	0.1669	0.1739	0.399	0.4063	0.3975	0.3962	0.6503	0.6516	0.6531	0.6486
	LSSH	0.2131	0.2177	0.2144	0.228	0.3887	0.3848	0.3874	0.3896	0.6392	0.6644	0.6586	0.6682
	DCH	0.3192	0.3377	0.3535	0.3722	0.5595	0.5793	0.5944	0.6015	0.7176	0.724	0.7285	0.7314
	FSH	0.2357	0.251	0.2524	0.2513	0.3578	0.3826	0.3912	0.3994	0.6457	0.6594	0.6618	0.6712
	SCRACH	0.3428	0.3497	0.3662	0.3613	0.7065	0.7045	0.7161	0.7248	0.6072	0.6332	0.6456	0.6474
	GSPH	0.2392	0.2772	0.2965	0.3018	0.7159	0.7346	0.7413	0.7453	0.5489	0.5725	0.5832	0.5887
	BATCH	0.3553	0.3781	0.3837	0.3940	0.7318	**0.7408**	0.7416	0.7464	0.6274	0.6505	0.6632	0.6638
	SDDH	0.3544	0.3739	0.3857	0.3904	0.721	0.7394	**0.7454**	0.7494	0.6504	0.6564	**0.6670**	0.6633
	SLCH	0.3526	0.3648	0.3537	0.3829	0.6294	0.6238	0.6269	0.6472	0.6161	0.6036	0.6367	0.6425
	ASCMH	**0.3664**	**0.3948**	**0.3911**	**0.3950**	**0.7319**	0.7371	0.7367	**0.7523**	**0.6510**	**0.6588**	0.6633	**0.6638**
Text to Img	CMFH	0.4563	0.4751	0.4986	0.5299	0.4255	0.4278	0.4215	0.4187	0.3498	0.3435	0.3486	0.3529
	LSSH	0.494	0.5072	0.5288	0.5402	0.4101	0.413	0.4176	0.4151	0.4219	0.4228	0.4249	0.4192
	DCH	0.6724	0.7097	0.7216	0.7141	0.667	0.7083	0.7285	0.7390	0.7106	0.7103	0.7260	0.7223
	FSH	0.4092	0.4864	0.4961	0.5247	0.5869	0.5979	0.6186	0.6251	0.4334	0.4446	0.4714	0.4819
	SCRACH	0.7398	0.7538	0.7496	0.7513	0.7685	0.7773	0.7936	0.8007	0.7488	0.7747	0.7855	0.7893
	GSPH	0.2467	0.2690	0.2829	0.2823	0.7405	0.7567	0.7615	0.7705	0.6779	0.6893	0.6921	0.6962
	BATCH	0.7382	0.7448	0.7586	0.7502	0.8005	**0.8187**	0.8192	0.8256	0.7596	0.7782	0.7815	0.7837
	SDDH	0.7452	0.7554	0.7590	0.7618	0.7917	0.8132	0.8241	0.8328	0.7638	0.7790	**0.7945**	0.7980
	SLCH	0.7356	0.7251	0.7343	0.7354	0.6721	0.6682	0.6761	0.7064	0.7298	0.7281	0.7660	0.7678
	ASCMH	**0.7481**	**0.7607**	**0.7590**	**0.7628**	**0.8062**	0.8125	**0.8286**	**0.8339**	**0.7677**	**0.7824**	0.7898	**0.7990**

Experimental Result









The Top-N
precision curves

Conclusion

ASCMH presents a novel approach to address the challenges in large-scale multimedia retrieval. By learning a common latent space and leveraging semantic labels, ASCMH effectively constructs binary codes and optimizes similarity matrices to enhance search accuracy. The experimental results demonstrate the superiority of ASCMH over existing methods, highlighting its potential for practical applications in the field of multimedia retrieval. Further research and exploration of ASCMH can lead to advancements in cross-modal hashing techniques and contribute to the development of more efficient and accurate retrieval systems.

澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU

Thank You!