

令和 4 年 6 月 13 日現在

機関番号：17401

研究種目：基盤研究(B) (一般)

研究期間：2019～2021

課題番号：19H04140

研究課題名(和文) Perfect-codeのハッシングへの応用と数理解析

研究課題名(英文) Application of Perfect-code to hashing and mathematical analysis

研究代表者

上瀧 剛 (Koutaki, Gou)

熊本大学・大学院先端科学研究部(工)・准教授

研究者番号：20582935

交付決定額(研究期間全体)：(直接経費) 11,200,000円

研究成果の概要(和文)：SNSの普及により増大するマルチメディアデータ(画像・テキスト・動画)のデータ圧縮/検索の効率化の手段としてバイナリハッシングが注目されている。これまで申請者らはアダマール符号列を用いたハッシング法を考案し、その有効性を示してきた。これに対して、本研究ではより分離性の高いPerfect-Codeを用いたハッシングおよび複素数・4元数への拡張を行った。実際にWikiとMNISTを用いた評価実験において、提案する手法が従来のHC-SDHに比べてmAPが優位に向上できることを確認した。

研究成果の学術的意義や社会的意義

これまでPerfect-codeをバイナリハッシングに利用する試みはこれまでなく学術的な新規性がある。さらにそれが情報検索精度の向上に役立つことを示すことができた。また、これらの手法が複素数・4元数へ拡張できること、および性能改善につながることを示せたことは意義が高いと考える。

研究成果の概要(英文)：Binary hashing has attracted attention as a means of efficient data compression/retrieval of multimedia data (images, text and video), which is increasing with the spread of SNS. Until now, the applicants have devised a hashing method using Hadamard code sequences and have shown its effectiveness. In contrast, in this study, hashing using Perfect-Code, which is more separable, and extensions to complex and quaternionic numbers were carried out. In fact, evaluation experiments using Wiki and MNIST have confirmed that the proposed method can improve the mAP significantly compared to the conventional HC-SDH.

研究分野：画像処理

キーワード：バイナリハッシング アダマール行列 Perfect-Code クロスモーダル

1 . 研究開始当初の背景

スマートフォンの普及に伴い、Facebook、twitter、Instagramなどのソーシャルネットワークサービスの利用が増加し、膨大な量の画像、テキスト、動画、音声などのマルチメディアデータが日々アップロードされている。これらの大規模データから必要なデータを検索する、優先順位をつける、あるいは、ストレージに格納されたデータ量を削減する手段としてバイナリハッシングを利用できる。バイナリハッシングは、これらのマルチメディアデータ、あるいは、そのデータの特徴を表す実数の数値列である特徴ベクトル(元データに対するタグと言える)を、更にコンパクトな0と1の二値の数値列であるバイナリコードへと変換する。バイナリコード同士の類似性の比較は、ビット演算程度の簡単な計算で行なえ、検索に用いれば、検索時間を大幅に短縮できる。また、検索タグのデータ量を大幅に削減できる。従来法の中でも、Shenらが提案したSDH (Supervised Discrete Hashing) 法 [Shen+, CVPR2015]は、バイナリハッシングを混合整数計画問題として定式化し、簡単なモデルと解法ではあるが優れた性能を有する。SDHは継続して改良されており、線形緩和を用いた方法 [Gui+, TNNLS2016]や回帰を用いた Fast SDH 法 [Gui+, TPAMI2017]などが提案されている。本申請課題に至る経緯は、このSDHの改良が出発点である。

L	$d=4$	$d=6$	$d=8$	$d=10$	$d=12$	$d=14$	$d=16$
6	4	2	1	1	1	1	1
7	8	2	1	1	1	1	1
8	16	2	2	1	1	1	1
9	20	4	2	1	1	1	1
10	40	6	2	2	1	1	1
16	2048	256	32	4	2	2	2
17	2816-3276	256-340	36	6	2	2	2
18	5632-6552	512-673	64-68	10	4	2	2
19	10496-13104	1024-1237	128-131	20	4	2	2
20	20480-26168	2048-2279	256	40	6	2	2

図1 . Perfect-code の存在数 (20 ビットまで)

2 . 研究の目的

本研究の目的は、Perfect-code のバイナリハッシングへの応用による大規模マルチメディアデータに対する高速かつ高精度なクロスモーダル検索手法の確立である。しかし、Perfect-code を求めることは容易ではない。一般に Perfect-code を求めるためにはビット列の組合せを列挙して探すしか方法がなく、NP 困難な問題である。例えば、図1の表の $L=17$ および $d=6$ の Perfect-code は、256 個から 340 個存在する程度にしか分かっておらず、その具体的なビットコード列は現時点では未知である。以上のように、一般の Perfect-code の探索は困難であるが、次節で述べるように一部の Perfect-code や、実用レベルの近似解を我々が用いている定式化により見つけられる可能性を持っている。更に、Perfect-code をバイナリハッシングに応用するという考えはこれまでになく、また、これからもメディアデータが増加していくことから、学術・応用ともに重要であると考える。

3 . 研究の方法

Perfect-code を用いたハッシング

Perfect-code の存在数は現時点では $L=28$ 程度までしか分かっていない。すなわち、画像検索で頻繁に用いられる $L=32$ ビットにおいてすら、その Perfect-code は未知である。このような無制約時の Perfect-code の発見は、もはや申請者らの力では無理であるが、既に個数や個数の範囲(下界・上界)、互いのハミング距離が分かっている Perfect-code に関しては、枝刈探索などのアルゴリズムを用いて Perfect-code を求められる可能性がある。より長いビット長では Perfect-code を求めることが困難となることが予想される。その場合、Perfect-code の条件を緩めた Quasi-Perfect-code の利用を検討する。Quasi-Perfect-code を求めることは十分可能であり、それらをバイナリハッシングへの応用を検討する。

複素アダマール行列を用いたクロスモーダルハッシング

アダマール行列を複素数に拡張した複素アダマール行列が知られており、主に通信分野で使われている。本研究では複素アダマール行列のクロスモーダルハッシングへの応用を検討する。例えば、画像およびテキストの2種類のドメインが与えられた場合に、画像を実軸、テキストを虚軸に取り、SDH によるハッシングを適用する。現状のクロスモーダルハッシングでは2種類のドメインが殆どであり、複素アダマール行列で事足りると考えているが、複素アダマール行列を4元数に拡張したものもあり、4種類のドメインまで対応可能であると考えている。バイナリハッシングに複素数を導入した例はなく、その効果に興味がある。

4 . 研究成果

Perfect-code を用いたハッシング

Perfect-code を厳密に求めることは難しいため、木探索を用いて近似的な Perfect-code (疑似 Perfect-code ; Quasi-Perfect-code) を求めた。具体的には、まず求めるコードの数およびハミング距離の下界を定めてしまい、その条件を満たす符号化列を求める。木探索において、この条件を満たさない符号化列は直ちに分かるため、それ以降の探索を打ち切る。必要な個数の符号化列が求まらない場合は、木をトレースバックして次のノードを検索する。

図4 . 1 に、4 ビットで互いのハミング距離が2となる符号化列を4つ求める木探索の例を示す。この場合、0000、0011、0101、0110 が求まる。

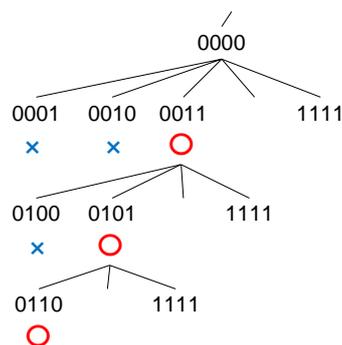


図4 . 1 木探索による Quasi-Perfect-code を求める例

上記アルゴリズムで32ビットの Quasi-Perfect-code を10個求めた結果例を図4 . 2 に示す。ハミング距離が18の符号を8つ生成できており、これはアダマール符号列のハミング距離16よ

りも符号間のハミング距離が大きい。この符号化列をハッシングに用いることでクラス間の距離が従来のアダマール符号列よりも大きくなり分類性能の向上が期待される。

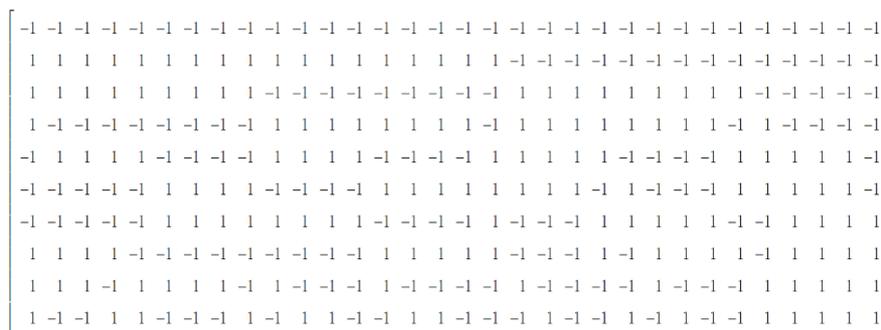


図4.2 生成した Quasi-Perfect-code の例 (32bit)

32 ビット以上の長い符号化列の場合は、アダマール符号列と同様に生成した 32bit 符号化列を連結することで 64-bit, 96-bit, 128-bit と拡張することができる。

上記の方法を用いて MS-COCO (サンプル数 328,000) および wiki (サンプル数 2,866) のクロスモーダルハッシングデータ (テキストと画像) に対して評価実験を行った。その時の Text2Img (テキストから画像の検索) の mAP (mean Average Precision) の結果を図 4.3 に示す。同図より、従来のアダマール符号列を用いた場合よりも提案する Perfect-code を用いた結果がより高い検索精度となることが確認できた。本研究成果は[1]の研究会が発表した。

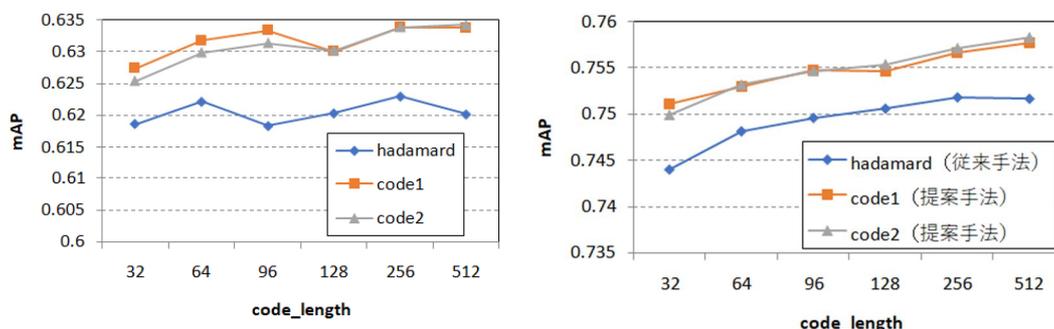


図4.3 Text2Img の評価結果 (左:MS-COCO, 右:wiki)

[1] 増田, 上瀧, “疑似 Perfect code を用いたクロスモーダルハッシングの検討”, PRMU 研究会, 2019.

複素アダマール行列を用いたクロスモーダルハッシング

バイナリハッシングに用いられる Kernel 特徴変換を複素数および 4 元数へ拡張する方式を考案した。すなわち、これまでの方法では学習サンプルをランダムに選択し (選択されたサンプルをアンカーと呼ぶ) それとの距離を用いた特徴変換が行われていた (図 4.4 (a)) が、そのアンカーを 2 個、ないし 3 個用いて変換する。これにより、従来の特徴変換ではサンプル間の距離のみで判別していたが、サンプル間の内積 (角度) を考慮した特徴変換が可能となる。アンカーを 2 個用いる場合は変換後の特徴ベクトルは複素数領域となり、3 個用いる場合は 4 元数領域となる。

このように、従来の実数のみに基づくアダマール符号列を用いたハッシングを R-HCSDH、今回の複素数および 4 元数を用いたハッシングを C-HCSDH および H-HCSDH と呼ぶことにする。特徴ベクトルが複素数や 4 元数になる場合でも、それを構成するハッシュ関数は R-HCSDH と同様に

アダマール行列で与えられることを数学的に証明した[2][3]。

図4.5に本手法を用いてwiki データセットでの検索性能評価を行った結果を示す。すべてのケースで提案する C-HCSDH および H-HCSDH が従来手法を上回る性能を得ることが確認され、提案法の有効性が確認できた。

[2] 末安, 京地, 上瀧, “四元数領域におけるアダマール符号離散ハッシング”, SIP 研究会, 2022.

[3] M. Sueyasu, S. Kyochi, G. Koutaki, “HADAMARD-CODED SUPERVISED DISCRETE HASHING ON COMPLEX AND QUATERNION DOMAIN”, IEEE ICIP (採択の見込み), 2022.

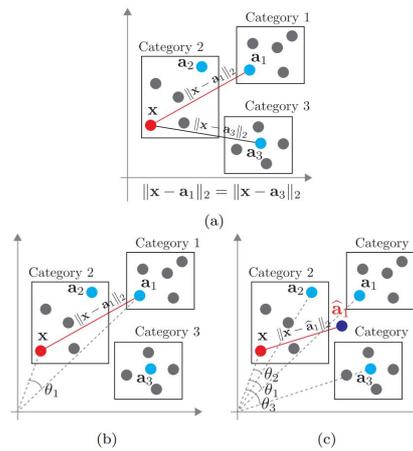


図4.4 複数アンカーを用いた Kernel 特徴変換

Wiki (Img2Img)				
L	SDH	\mathbb{R} -HCSDH	\mathbb{C} -HCSDH	\mathbb{H} -HCSDH
16	0.2352	0.2466	0.2517	0.2616
32	0.2509	0.2590	0.2593	0.2741
64	0.2617	0.2661	0.2646	0.2810
128	0.2651	0.2699	0.2676	0.2850
256	0.2400	0.2721	0.2692	0.2870
512	0.1858	0.2728	0.2706	0.2895

Wiki (Img2Txt)				
L	SDH	\mathbb{R} -HCSDH	\mathbb{C} -HCSDH	\mathbb{H} -HCSDH
16	0.2933	0.3059	0.3155	0.3144
32	0.3091	0.3218	0.3292	0.3308
64	0.3197	0.3294	0.3364	0.3390
128	0.3241	0.3339	0.3414	0.3438
256	0.3045	0.3365	0.3441	0.3460
512	0.2495	0.3370	0.3454	0.3490

Wiki (Txt2Img)				
L	SDH	\mathbb{R} -HCSDH	\mathbb{C} -HCSDH	\mathbb{H} -HCSDH
16	0.5511	0.5675	0.5761	0.5974
32	0.5857	0.5849	0.5909	0.6180
64	0.6021	0.5929	0.5979	0.6274
128	0.6093	0.5970	0.6014	0.6322
256	0.5867	0.5998	0.6043	0.6360
512	0.4785	0.5999	0.6051	0.6371

図4.5 wiki での評価結果

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計2件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 増田, 上瀧
2. 発表標題 疑似Perfect code を用いたクロスモーダルハッシングの検討
3. 学会等名 PRMU研究会
4. 発表年 2019年

1. 発表者名 未安, 京地, 上瀧
2. 発表標題 四元数領域におけるアダマール符号離散ハッシング
3. 学会等名 SIP研究会
4. 発表年 2022年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	白井 啓一郎 (Shirai Keiichiro) (00447723)	信州大学・学術研究院工学系・准教授 (13601)	
研究分担者	京地 清介 (Kyochi Seisuke) (70634616)	工学院大学・情報学部(情報工学部)・准教授 (32613)	

6. 研究組織（つづき）

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 分 担 者	松井 勇佑 (Matsui Yusuke) (80780676)	東京大学・大学院情報理工学系研究科・講師 (12601)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関