

令和元年6月13日現在

機関番号：12608

研究種目：基盤研究(C) (一般)

研究期間：2015～2018

課題番号：15K06055

研究課題名(和文)画像マッチングのための大型画像変換に関する研究

研究課題名(英文)Study on Transform with Large Size for Image Matching

研究代表者

伊藤 泉 (Ito, Izumi)

東京工業大学・工学院・助教

研究者番号：90583435

交付決定額(研究期間全体)：(直接経費) 3,300,000円

研究成果の概要(和文)：本研究では、計算コストおよびメモリ使用量を大幅に削減できる、画像マッチング機能を有する大型画像変換の設計をおこなった。疎行列分解に基づく大型近似離散コサイン変換の設計、および任意サイズハール変換の設計とそれらの評価を行った。大型近似離散コサイン変換では、計算アルゴリズムを単純化し、浮動小数点の乗算を整数の乗算とシフト演算のみでおこなうことにより、従来法と同程度の画像マッチング精度を維持しつつも、計算コストを大幅に削減した。任意サイズハール変換では、積分画像を効率よく利用できるように設計した基底生成関数を導出し、高精度の全探索法と等価でありながらも、計算コストを大幅に削減した。

研究成果の学術的意義や社会的意義

これまで、ハードウェア構成を考慮した離散コサイン変換は、主に8x8のサイズに限定され、ハール変換は適用サイズが2の冪乗に限られていた。画像の解像度が飛躍的に増大する今日では、より大きなサイズに適応した、計算コストおよびメモリ使用量を抑制した変換が必要となる。今回開発した近似コサイン変換、および任意サイズハール変換は、高い画像マッチング機能を有しながらも、計算量を大幅に削減可能であり、安価なデバイスにおいても使用可能である。

研究成果の概要(英文)：Discrete cosine transform approximation with large size and Haar transform with arbitrary size were designed for image matching and their accuracy in image matching were evaluated. In the discrete cosine transform approximation based on matrix factorization algorithms, the computational complexity is reduced due to the architecture simplified by substituting floating-point multipliers with integer multipliers and shift operators, and the accuracy of image matching is almost the same as the conventional one. In the Haar transform with arbitrary size, the mathematical expression for basis images was derived designing for the efficient use of an integral image. Keeping the accuracy of the full search method, the transform reduces the computational load drastically.

研究分野：信号処理，画像処理

キーワード：画像マッチング 離散コサイン変換 ハール変換

様式 C-19、F-19-1、Z-19、CK-19 (共通)

1. 研究開始当初の背景

画像マッチング技術は、画像・映像信号処理やコンピュータビジョンにおける基本処理であり、二枚の画像の対応づけを行う処理である。その応用は、画像検索、動き推定、オブジェクト抽出、バイオメトリクス認証、超解像度処理等、広範囲にわたる。中でも、位相限定相関法は、高精度な位置ずれ推定をおこなう画像マッチング手法で知られている。離散コサイン変換 (Discrete Cosine Transform: DCT) は、エネルギー凝縮能力に優れ、JPEG、MPEG、H.26x 等の標準画像圧縮方式において用いられている。これらの方式では、画像はブロックと呼ばれる小領域 (主に 8×8 画素) に分割され、ブロックごとに DCT が適用される。浮動小数点演算を必要とする任意サイズの高速 DCT アルゴリズムは提案されているが、ハードウェア実現を考慮した演算アルゴリズムは、圧縮方式の要請から、主に 8×8 画素用 DCT に限られている。画像マッチング技術は古くから研究されているが、基本性能として、照明変化やノイズに対するロバスト性、演算の高速性、および、マッチングの高精度性の追求のほか、実用化における要請から、パラメータの少なさや、省メモリであること、また、あらゆる画像 (画像サイズ、テクスチャの量の違いなど) に対応できる汎用性が求められ、いまなお研究が進められている。

2. 研究の目的

本研究では、申請者が提案した画像マッチング手法の基礎の確立として、計算コスト、およびメモリ使用量を大幅に削減できる、大型画像変換の設計をおこない、応用展開として高解像度画像のための新たな画像マッチング・圧縮処理技術を構築する。既存の浮動小数点演算を必要とする高速アルゴリズムに基づき、近似誤差を抑え、乗算器数およびメモリ使用量を削減する、ハードウェア構成を考慮した大型近似離散コサイン変換 (大型近似 DCT) の設計手法を明らかにする。また、DCT に限らず、計算コスト、メモリ使用量を大幅に削減可能な画像マッチング機能付き大型変換を構築することを目的とする。

3. 研究の方法

3.1 疎行列分解による高速アルゴリズムに基づく大型近似 DCT の設計とその評価

疎行列分解に基づく浮動小数点演算を必要とする種々の高速 DCT アルゴリズムをもとに、乗算器構成を単純化し、各乗算器係数を一定語長の値で近似することにより、演算量およびメモリ使用量を大幅に削減する大型近似 DCT を作成した。また、乗算器係数ごとに異なる有限語長を与えたとき、全体の近似精度に大きな差を生じうる乗算器を特定するため、多様な画像データベースを用いてシミュレーションをおこない、各乗算器係数の寄与率を検討した。MATLAB 上で実装した大型近似 DCT を用いて画像マッチングをおこない、マッチング精度を比較した。

3.2 他の離散変換による可変大型 DCT 変換の設計とその評価

Sylvester アダマール変換および Paley アダマール変換を用いて、小さなブロックサイズの DCT 係数から大きなブロックサイズの DCT 係数に変換し、大型近似 DCT を生成することを試みた。この可変大型 DCT を MATLAB 上で実装し、その計算コスト、およびメモリ使用量を調査した。

3.3 任意サイズハール変換の設計とその評価

ハール変換は、Fourier 変換や DCT と異なり、変換のための乗算係数が全て異なるわけではないため、積分画像を利用すれば、数回の加減算のみで総和が求められ、計算コストを飛躍的に削減することが可能となる。しかしながら、適用できるサイズは 2 のべき乗サイズに限られていた。このハール変換の計算コストに着目し、相關計算によるマッチングとは異なる、ブロックマッチングを検討し、2 のべき乗以外のサイズに適用できる、積分画像に適した任意サイズハール変換を設計した。MATLAB で実装し、シミュレーションにより画像マッチング性能、および、処理時間を計測した。また、変換によるエネルギー凝縮能力の観点、および、エントロピーから、画像データベースを用いて最良基底画像選択アルゴリズムを検討した。さらに、既存のノイズ除去法 (ノンローカルミーン法、特異値分解法) との組み合わせにより、その演算の高速性、およびノイズ下におけるマッチングのロバスト性を評価した。

4. 研究成果

4.1 疎行列分解による高速アルゴリズムに基づく大型近似 DCT の設計

任意サイズの高速 DCT アルゴリズムをもとに、乗算器構成を単純化し、乗算器係数を一定語長の値で近似することにより、演算量およびメモリ使用量を大幅に削減できる大型近似 DCT を開発した。図 1 に、Wang の高速アルゴリズム (素因数分解することにより、再帰的に DCT を実行) にもとづく、 $N=16$ の DCT のシグナルフロー図を示す。信号は左端から入力され、右方向へ順次、乗算器 (\times , C , および S) および加算器 (\oplus) により計算される。乗算器はそれぞれ、 $\times=1/\sqrt{2}$, $\times=\sqrt{2}$, $C_{i/j} = \cos(i\pi/j)$, および $S_{i/j} = \sin(i\pi/j)$ を表している。この乗算器の組み合わせを、図 2 に示すように、二つのリフティング構成 (p, u, p' , および u' は乗算係数) を用いて単純化した。この結果、 N 点 DCT に対して、乗算数が $(N/2)(\log_2 N - 1)$, 加減算数が $(3N/2)(\log_2 N - 2) - N/2 + 6$ となった。つぎに乗算器係数を整数乗算 Q とシフト演算のみ 2^E (E は整数) で演算できるように、乗算器係数を $Q \times 2^E$ と近似した。

表 1 に、自己回帰モデル (AR(1))における、DCT 係数と近似 DCT 係数の平均二乗誤差を示す。ブロックサイズ ($N \times N$) が大きくなるほど誤差が大きくなるが、 E (右シフト数) を大きくすることで、近似誤差を小さくできることがわかる。次に画像 (大きさ 256×256) に近似 DCT を適用したときの正負符号が、従来の DCT 正負符号と異なる割合 (パーセンテージ) を種々の画像データベースを用いて比較した。平均で、 $E = 1$ のときは 18%, $E=2$ のときは 12%, $E=3$ のときは 6.3% となり、 $E=4$ 以降は E が 1 大きくなるにつれて、そのパーセンテージは半減した。変換係数の正負符号が異なる場所については、集中して異なる箇所はなく、変換係数全体に異なる符号が点在していた。画像によりその場所も異なった。図 3 に、この近似 DCT の正負符号を用いておこなう画像マッチング例を示す。この例では二枚の画像は、縦横に 100 ピクセルずつ平行移動した関係にある。 $E=1$ のとき、正負符号の違いが、符号位相相関関数においてやや雑音として現れるが、1 ピクセル毎の平行移動量推定は可能であった。 $E=2$ 以降では、従来の DCT を用いた画像マッチングと同程度の精度を有することを確認した。

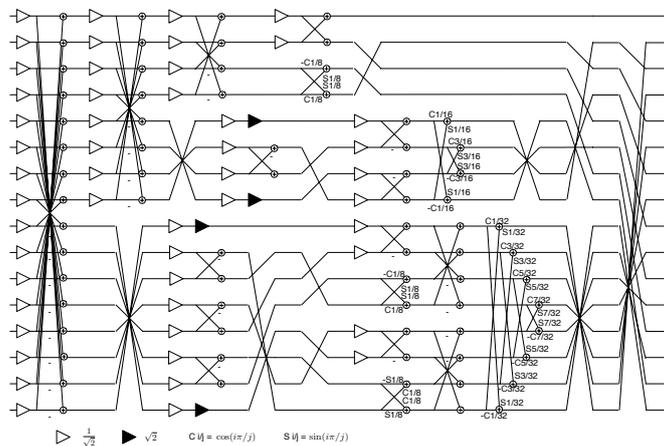


図 1 疎行列分解による DCT の構成 ($N=16$)

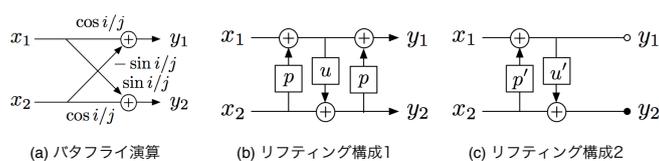
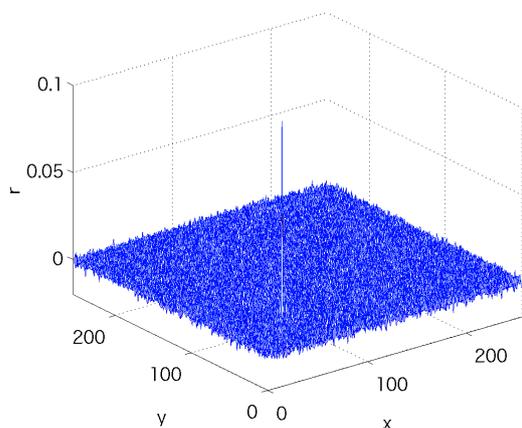


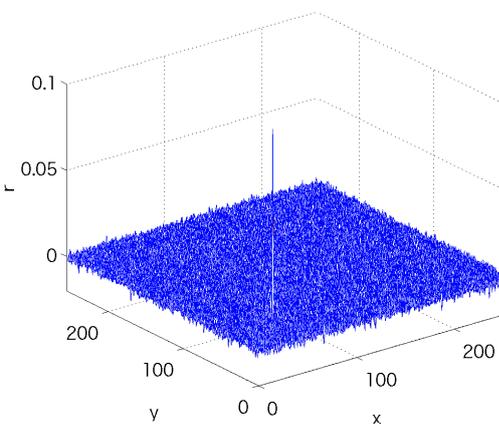
図 2 乗算器の単純化

表 1. AR(1)モデルにおける DCT 係数と近似 DCT 係数の平均二乗誤差

| N | E | | | | | | | |
|-----|--------|----------|----------|----------|----------|----------|----------|----------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 8 | 0.004 | 9.30E-04 | 3.94E-04 | 2.33E-05 | 1.31E-05 | 7.36E-06 | 2.93E-07 | 1.43E-07 |
| 16 | 0.0064 | 0.0035 | 5.58E-04 | 1.46E-04 | 3.44E-05 | 1.27E-05 | 2.15E-06 | 4.69E-07 |
| 32 | 0.0135 | 0.0051 | 0.0015 | 1.97E-04 | 8.68E-05 | 2.35E-05 | 3.08E-06 | 1.32E-06 |
| 64 | 0.0211 | 0.0089 | 0.0021 | 4.53E-04 | 1.27E-04 | 4.37E-05 | 5.55E-06 | 1.89E-06 |
| 128 | 0.0326 | 0.0125 | 0.0032 | 6.39E-04 | 2.03E-04 | 6.17E-05 | 9.65E-06 | 2.99E-06 |
| 256 | 0.0453 | 0.0167 | 0.004 | 8.81E-04 | 2.62E-04 | 7.86E-05 | 1.26E-05 | 4.02E-06 |



(a) 従来の DCT



(b) 近似 DCT ($E=2$)

図 3 符号位相相関関数例

4.2 ブロック DCT の併合による大型近似 DCT 変換の設計

アダマール変換は、種々の変換タイプがあり、それらはグレイコード、バイナリコード、およびそれらのビット反転したコードを介して、タイプ変換が可能である。これを利用し、容易に小さなサイズの変換係数を併合することにより、大きなサイズの変換係数を生成することが可能となるシルベスタアダマール変換をウォルシュアダマール変換にタイプ変換し、DCT を生成した。しかしながら、その演算コスト、および近似誤差は、疎行列分解による高速アルゴリズムに基づく大型近似 DCT に比べて劣っていた。

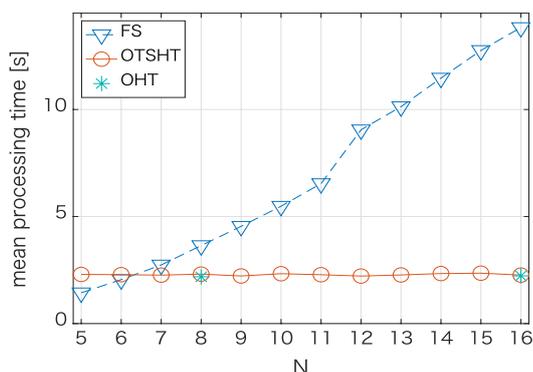


図 4 ブロックマッチング平均実行時間

4.3 任意サイズハール変換の開発

2 のべき乗以外のサイズに適用できる、任意サイズ二次元ハール変換関数を導出した。積分画像を効率よく利用するため、一つの基底画像に最大二つの一定数の領域を有するように設計した。

いま、画像の中に、あるクエリー画像 (サイズ $N \times N$) と同一あるいは類似した部分を探査することを考える。全探索法は画像内のすべての $N \times N$ の領域の画素値をクエリー画像と比較するため、マッチング精度は最も高いが、計算コストが高い。ハール変換を用いた画像マッチングは、ハール変換が正規直交変換であることから、全探索法の精度を維持しつつも、

誤差を計算する領域を絞り込むことができる。高速フーリエ変換を用いて計算した相関を利用する全探索法と等価な高速手法も提案されているが、それと比較しても、ハール変換を用いる手法は高速である。ベンチマーク画像を用いて、全探索法 (FS)、従来のハール変換による探索法 (OHT)、および、任意サイズハール変換による探索法 (OTSHT) のクエリー画像サイズ ($N \times N$) の違いによる平均実行時間を図 4 に示す。ここで画像サイズは 512×512 である。クエリーサイズが大きくなるにつれ、全探索法では処理時間が増大するのに対して、任意サイズハール変換は、処理時間が一定に保持されることがわかる。

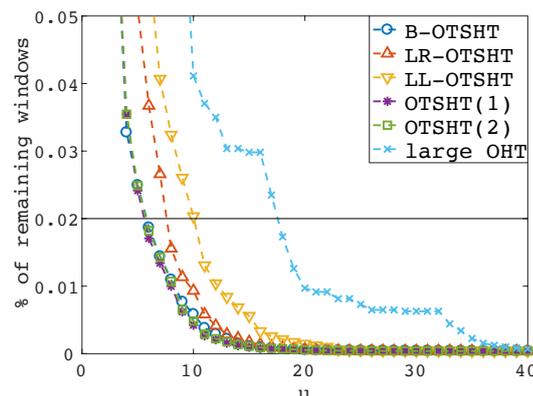


図 5 候補ウィンドウの絞り込み性能例, $N=9$

開発した任意サイズハール変換は、二分木構造から生成し、その木構造により変換基底が異なる。 $N+1$ 枚の葉を持つ二分木の総数は、カトラン数で表され、 N が大きくなるにつれて膨大な数となる。例として、 N が 12 のとき、58,786 もの異なる変換基底セットが存在する。エントロピーをはじめ種々の尺度による最良基底の選択を試みたが、画像およびクエリー画像に依存するため、一般に、平衡木が良い結果を出すことのみを示した。図 5 に、例として、クエリーサイズが 9×9 のときの異なる基底画像セットにおける基底画像数 u の違いによる残存候補ウィンドウ数の全体の割合 (パーセンテージ) を示す。ここでは、代表的な二分木構造である平衡木 (B-OTSHT)、右対数木 (LR-OTSHT) および左対数木 (LL-OTSHT)、二つの任意の二分木 (OTSHT(1) および OTSHT(2))、および $N=16$ の従来のハール変換 (Large OHT) を比較している。平衡木は、基底画像数 $u=6$ のとき、全体の 0.02% 以下まで、探索候補を絞り込んでいることがわかる。また、従来の 2 のべき乗のハール変換では、候補の絞り込みは、任意サイズハール変換と比べて劣ることがわかる。

さらに、開発した任意サイズハール変換を用いた応用例として、ノイズ除去法に組み込むことにより、ノイズ下におけるマッチング精度を示した。ノイズ除去法は、全探索法の代わりに、任意サイズハール変換を用いて類似ブロックを集め、この類似ブロックから元画像を推定する。一枚の画像内からの類似ブロックの探索のほか、複数枚の画像内からの類似ブロック探索を構築した。現在のノイズ除去法のベンチマークでもある BM3D 法をはじめ、種々の最先端のノイズ除去法との比較実験をおこなった。マッチング精度の最も高いといわれる全探索法を用いた BM3D 法に比べて、ピーク信号対雑音比において平均約 0.3 [dB] 上回る結果を得た。

5. 主な発表論文等

[雑誌論文] (計 1 件)

1. I. Ito, K. Egiazarian, Two-dimensional orthonormal tree structured Haar transform for Fast block matching, MDPI Journal of Imaging, 4(11), 131, 2018 ,

[学会発表] (計 7 件)

1. I. Ito, K. Egiazarian, Full Search Equivalent Fast Block Matching using Orthonormal Tree-structured Haar Transform, International Symposium on Image and Signal Processing Analysis, pp.177-182. Sep. 2017
2. I. Ito, K. Egiazarian, Design of orthonormal Haar features for fast block matching, EURASIP European Signal Processing Conference, ISBN 978-0-9928626-7-1, pp.2452-2456, Sep. 2017
3. 伊藤 泉, 画像マッチングのための任意サイズ Haar 変換の設計についての検討, 電子情報通信学会 信学技報, pp.14-17, Aug. 2018
4. 伊藤 泉, 任意サイズ Haar 変換を用いた SVD に基づく画像ノイズ除去, 電子情報通信学会総合大会, 情報・システム講演論文集 2, p. 31, 2018
5. 伊藤 泉, 任意サイズ Haar 変換を用いた画像ノイズ除去の検討, 電子情報通信学会 信号処理シンポジウム pp.427-432, 2017
6. 伊藤 泉, 任意サイズ Haar 変換と画像マッチングへの応用, 電子情報通信学会 技術研究報告, vol.117, no. 96, pp.87-91, Jun. 2017
7. 伊藤 泉, Heikki Huttunen, Karen Egiazarian, 任意サイズ正規直交 Haar 変換による高速ブロックマッチング, 電子情報通信学会 総合大会情報・システム講演論文集 2, p.25 Mar. 2017

[図書] (計 0 件)

[産業財産権]

○出願状況 (計 0 件)

名称：
発明者：
権利者：
種類：
番号：
出願年：
国内外の別：

○取得状況 (計 0 件)

名称：
発明者：
権利者：
種類：
番号：
取得年：
国内外の別：

[その他]

ホームページ等

6. 研究組織

(1)研究分担者

研究分担者氏名：

ローマ字氏名：

所属研究機関名：

部局名：

職名：

研究者番号 (8 桁)：

(2)研究協力者

研究協力者氏名：Karen Egiazarian

ローマ字氏名：Karen Egiazarian

※科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。