

量子情報理論に基づく Karhunen-Lòeve 変換の情報集約性に関する理論的考察 Analytical evaluation on the energy compaction provided by Karhunen-Lòeve transform based on quantum information theory

坂東幸浩², 大林浩気¹, 高村誠之², 如澤裕尚²,

Yukihiro BANDO², Hiroki OHBAYASHI¹, Seishi TAKAMURA², and Hirohisa JOZAWA²

1 まえがき

画像符号化における重要な要素技術の一つに、離散コサイン変換(DCT)に代表される変換符号化がある。画像符号化における変換符号化の役割は、空間的な画素間相関の除去を行うことにある。変換符号化により少数の変換係数に情報を集中させ、量子化により情報の集中度の低い変換係数を切り捨てることで、符号化対象信号に対する情報量が削減される。従って、変換符号化に対する符号化効率、変換係数の情報の偏在度で評価することができる。つまり、少数の変換係数に情報を集中させることができる変換符号化は、符号化効率の高い変換符号化と言える。

変換符号化の符号化効率は、変換の対象となる信号の空間相関と変換符号化の次元数に依存する。例えば、H.264/AVC FRExt[1]における検討において、HDTVサイズの映像信号に対しては、 4×4 次元のDCTに加えて、 8×8 次元のDCTを用いる事で、符号化効率が向上する事が報告されている。さらに、HDTVを超える解像度の映像信号に対しては、 4×4 次元および 8×8 次元のDCTとあわせて、 16×16 次元のDCTを併用することで符号化効率が向上することが報告されている[2]。このように、空間相関の高い画像に対しては、高次元の変換符号化を用いる事で、符号化効率が向上することが知られている。

一方、変換符号化の次元数と符号化効率の関係に対して、理論的な根拠を与えるための解析は十分なされていない。Somanら[3]は、直交変換がパラユニタリフィルタバンクの一種であることを利用して、パラユニタリフィルタバンクとして一般化した形式で、直交変換の符号化性能を coding gain[4]に基づいて議論している。しかし、変換符号化の効率と次元数の関係については検討されていない。Andrewら[5]は coding gain とサブバンドフィルタのタップ長の関係を調査している。しかし、タップ長の異なるフィルタを用いて比較実験の結果を検証した定性的な説明にとどまり、両者の関係について定量的な説明はなされていない。Vaidyanathan[6]は、coding gain を Karhunen-Lòeve 変換(KL変換)の固有値を用いてモデル化し、coding gain と KL 変換の次元数との関係を議論している。このモデルは、次元数の増加に伴う coding gain の単調増加性は説明して

いるが、その増加率の変化については説明できていない。これは、入力信号の画素間相関が、coding gain に及ぼす影響について考察できていないためである。こうした増加率は入力信号の画素間相関と密接に関係しているため、増加率まで含めたモデル化を行うには、入力信号の画素間相関が coding gain に及ぼす影響を考察する必要がある。

変換符号化の次元数を固定した場合、変換後の情報の偏在度に関して最適な変換基底を与えるのは KL 変換である。従って、変換符号化の次元数と変換効率の関係を考察する対象として、KL 変換を用いるのは自然な選択といえる。そこで、KL 変換について次元数と変換効率の関係をモデル化するため課題を整理する。KL 変換後の情報は、固有ベクトル・固有値として表現されており、変換後の情報の偏在度(情報集約度)は、固有値という形に集約される。従って、そうした情報集約度と KL 変換の次元数との関係を議論するためには、固有値という形に凝縮された情報に対して、その内部に潜む KL 変換の次元数と入力信号の関係を観測可能な形で定式化する必要がある。しかし、KL 変換後の状態では、定量的に評価したい情報間の関係が直接的には観測できず、観測可能な情報は固有値に限定されるという点に、問題の難しさがある。

そこで、本稿では、この問題を解決するための数学的道具として、量子情報理論における密度作用素と von Neumann エントロピーを用いる。密度作用素は確率分布の量子力学的な拡張であり、von Neumann エントロピーは確率分布に関する Shannon エントロピーの量子力学的な拡張である。これらの数学的道具を用いることで、KL 変換の次元数と情報集約度の関係の評価を量子力学的な確率系に対する議論に帰着することができ、複合系から部分系を取り出す操作を通して、次元数の異なる KL 変換の情報集約度を定量的に評価することが可能となる。さらに、入力信号の空間的な画素間相関の影響も自然な形で考慮に加えることができることから、KL 変換の次元数と入力信号の空間的な画素間相関が符号化効率に及ぼす影響を解析的にモデル化し、KL 変換の符号化効率を定量的に評価すること事が可能となる。

¹日本電信電話株式会社 NTT サイバースペース研究所

²九州大学大学院 数理学府

2 KL変換の情報集約度の定量的評価

2.1 KL変換

KL変換について、表現法を整理しておく。 d 次元実ベクトル空間 \mathbb{R}^d の n 個のベクトル $\xi_i = (y_{i1}, \dots, y_{id}) \in \mathbb{R}^d$ ($i = 1, \dots, n$) が与えられているとする。各 ξ_i の第 k 成分を並べて得られる n 次元列ベクトルを y_k とおく。すなわち、

$$y_k = \begin{pmatrix} y_{1k} \\ \vdots \\ y_{nk} \end{pmatrix} \in \mathbb{R}^n \quad (k = 1, \dots, d)$$

とおく。

$$V = \frac{1}{n} \sum_{i=1}^n (\xi_i - m)^t (\xi_i - m)$$

により定まる $d \times d$ 正方行列 V を y_1, \dots, y_d の分散共分散行列と呼ぶ。ここで、 $m = \frac{1}{n} \sum_{i=1}^n \xi_i$ とした。全ての要素を $\frac{1}{n}$ とする $n \times n$ 行列を E_n とし、 $n \times n$ の単位行列を I_n とし、 $E_n^+ = I_n - E_n$ と表すことにする。これにより V は

$$V = \frac{1}{n} \left[(E_n^+ y_k)^t (E_n^+ y_l) \right]_{k,l=1,\dots,d}$$

と表すことができる。但し、 $[a_{kl}]_{k,l=1,\dots,d}$ は k, l 成分を a_{kl} とする $d \times d$ 行列を表す。また、 $(E_n^+ y_k)^t$ は行列 $(E_n^+ y_k)$ の転置行列を表す。 V は半正定値行列である。従って、固有値はすべて0以上の実数であり、固有ベクトルはそれぞれ互いに直交する。 V の固有値を $\lambda_1, \dots, \lambda_d$ ($\lambda_1 \geq \dots \geq \lambda_d \geq 0$)、それに対応するノルム1の固有ベクトルを $r_1, \dots, r_d \in \mathbb{R}^d$ とする。 r_1, \dots, r_d を変換基底とする直交変換を ξ_1, \dots, ξ_n に関するKL変換と呼ぶ。また、 $p_k = \lambda_k / \text{tr} V$ を r_k の寄与率と呼び、 $\sum_{i=1}^k p_i$ を r_1, \dots, r_k の累積寄与率と呼ぶ。但し、 $\text{tr} A$ は行列 A のトレースを表す。これはKL変換における変換係数の情報の r_k への集中度を表すものであり、 $p_1 \geq \dots \geq p_d \geq 0$, $\sum_{k=1}^d p_k = 1$ を満たす。以下、寄与率から得られる確率分布 $\psi = (p_1, \dots, p_d)$ を寄与率分布と呼ぶことにする。 ξ_1, \dots, ξ_n を縦に並べることで得られる $n \times d$ 行列を Y とする。即ち、

$$Y = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = \begin{pmatrix} y_{11} & \cdots & y_{1d} \\ \vdots & \ddots & \vdots \\ y_{n1} & \cdots & y_{nd} \end{pmatrix}$$

とおく。このとき、 $Y = (y_1, \dots, y_d)$ であるから、次のように表すことができる。

$$\begin{aligned} V &= \frac{1}{n} \begin{pmatrix} (E_n^+ y_1)^t \\ \vdots \\ (E_n^+ y_d)^t \end{pmatrix} (E_n^+ y_1, \dots, E_n^+ y_d) \\ &= \frac{1}{n} \begin{pmatrix} y_1^t \\ \vdots \\ y_d^t \end{pmatrix} E_n^{+*} E_n^+ (y_1, \dots, y_d) \\ &= \frac{1}{n} Y^t E_n^+ Y \end{aligned}$$

2.2 Kullback-Leibler divergence に基づく情報集約度の評価尺度

本節では、KL変換の符号化効率の定量的な評価について Shannon エントロピーの観点から議論を行い、KL変換の符号化効率を示す関数を導入する。

まず、信号に適用するKL変換の次元数が固定されている場合を考える。事象の個数 d を固定した場合、確率分布 $\psi = (p_1, \dots, p_d)$ の偏りは、その Shannon エントロピー $H(\psi) = -\sum_{k=1}^d p_k \log p_k$ が小さいほど大きいといえる。このことから、KL変換の次元数 d が固定されている場合、寄与率分布の Shannon エントロピーの大小によって、KL変換に対する符号化効率を評価することができる。例えば、2つの入力信号に対して、 d 次元のKL変換を適用し、寄与率分布が各々、 $\psi = (p_1, \dots, p_d)$, $\phi = (q_1, \dots, q_d)$ であったとする。このとき、 ψ と ϕ の Shannon エントロピー $H(\psi), H(\phi)$ に対して $H(\psi) < H(\phi)$ が成立するならば、 ϕ よりも ψ の方が、寄与率分布の偏りが大きく、符号化効率も高いといえる。

一方、信号に適用するKL変換の次元が固定されていない場合は、確率分布の事象数が異なってしまうため、寄与率分布の Shannon エントロピーの大小では、KL変換に対する符号化効率を評価することはできない。例えば、2つの入力信号に対して、一方には次元数4のKL変換を、他方には次元数8のKL変換を適用し、各々の寄与率分布が $\psi = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$, $\phi = (\frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8})$ であったとする。このとき、同程度の累積寄与率を達成するために必要な最小の基底ベクトル数の割合はどちらも一致していることがわかる。例えば、累積寄与率が $3/4$ に達するために必要な最小の基底ベクトル数はそれぞれ3.6であり、どちらも全体の次元数の $3/4$ となっている。従って、KL変換に対する符号化効率はどちらの入力信号も等しいと考えるべきである。ところが、それぞれの Shannon エントロピーを求めると、 $H(\psi) = 2 \log 2$, $H(\phi) = 3 \log 2$ となり一致していない。このような問題を回避するため、符号化効率を示す基準として、KL変換の次元数の違いに影響を受けない尺度

を新たに導入する必要がある。

そこで、Kullback-Leibler divergence [7] と呼ばれる情報量を利用する。二つの確率分布 $\psi = (p_1, \dots, p_d), \phi = (q_1, \dots, q_d)$ に対して

$$H(\psi|\phi) = \sum_{k=1}^d p_k (\log p_k - \log q_k)$$

によって定義される量を ϕ に対する ψ の Kullback-Leibler divergence と呼ぶ。これは二つの確率分布 ψ, ϕ の間の隔たりの大きさを表すものであり、常に $H(\psi|\phi) \geq 0$ となる。また、 $H(\psi|\phi) = 0$ となるのは $\psi = \phi$ のときに限る。以下、事象数 d の一様分布を u_d で表すとす。このとき、 u_d は d 個の要素が全て $\frac{1}{d}$ となるベクトルである。一様分布とは同事象数の確率分布の中で最も偏りの無いものであるから、一様分布からの隔たりが大きい確率分布ほど、つまり一様分布に対する Kullback-Leibler divergence の値が大きい確率分布ほど偏りの大きい確率分布であると考えられることができる。

この考え方にに基づき、符号化効率を表す関数として、一様分布 u_d に対する寄与率分布の Kullback-Leibler divergence

$$T(\psi) = H(\psi|u_d) = \log d - H(\psi)$$

を導入する。すなわち、 $T(\psi)$ の値が大きければ大きい程、符号化効率が良いと考える。この T は、符号化効率に関して次元数の違いが影響しない（次元数によらず常に符号化効率が最も悪い）一様分布を基準にした相対的な値として定義されているため、次元数の違いの影響を受けずに符号化効率を評価することができる。以下では、 $T(\psi)$ を ψ に対する情報集約度と呼ぶ。上で述べた例 ($\psi = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}), \phi = (\frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \dots, \frac{1}{8})$) において $T(\psi), T(\phi)$ を求めると $T(\psi) = T(\phi) = 0$ となり、確かに一致している。また、

$$\begin{aligned} T(\psi) > T(\phi) &\iff \log d - H(\psi) > \log d - H(\phi) \\ &\iff H(\psi) < H(\phi) \end{aligned}$$

であることから、本節前半で行った KL 変換の次元数 d が固定されている場合の議論とも矛盾しないことがわかる。

以下、情報集約度 T の定義の妥当性を示す根拠として、本節前半の議論との関連についてさらに述べる。与えられた2つの入力信号に対し、一方は次元数 d 、もう一方は次元数 $2d$ の KL 変換を適用した結果、寄与率分布がそれぞれ $\psi = (p_1, \dots, p_d), \phi = (q_1, \dots, q_{2d})$ であったとする。本節前半で議論した、次元数が固定されている場合の Shannon エントロピーによる符号化効率の評価法を応用することにより、これらの KL 変換の符号化効率の比較を行うことを考える。最も素朴な方法として、次元数が少ない方の寄与率分布をさらに細かく分割し、次元数をそろえるという方

法が考えられる。そこで $\psi = (p_1, \dots, p_d)$ の代わりに各 p_k を適当に2分割して得られる事象数 $2d$ の確率分布

$$\psi' = (p'_1, p'_2, \dots, p'_{2d}), p'_{2k-1} + p'_{2k} = p_k \quad (k=1, 2, \dots, d)$$

を考える。ただし、 $p'_1 \geq p'_2 \geq \dots \geq p'_{2d}$ を満たしているとする。このとき、累積寄与率と必要な最小基底ベクトル数の割合の関係は分割前と変わっておらず（分割前、分割後どちらも累積寄与率が $\sum_{k=1}^a p_k$ に達するために必要な最小基底ベクトル数の割合は a/d となっている）、したがって符号化効率は分割前と等しいと考えてよい。さらに、 ψ' と ϕ は事象数がどちらも $2d$ であるから、直接 Shannon エントロピー $H(\psi'), H(\phi)$ の値を比較しても問題ない。また、上の条件を満たす分割の方法は一意とは限らないが、ここでは各 p_k をちょうど $1/2$ に分割するとする。すなわち、 $\psi' = (p_1/2, p_1/2, p_2/2, p_2/2, \dots, p_d/2, p_d/2)$ とする。このとき、 ψ' は ψ と確率分布 $u_2 = (\frac{1}{2}, \frac{1}{2})$ の直積によって得られる確率分布となっており、従って Shannon エントロピーの理論における条件付エントロピーの計算から、

$$H(\psi') = H(\psi) + H(u_2) = H(\psi) + \log 2$$

が導かれる。以上のことをまとめると、これらの KL 変換の符号化効率を比較するには、 $H(\psi) + \log 2$ の値と $H(\phi)$ の値の大小を比較すればよいということになる。ここで、

$$\begin{aligned} H(\psi) + \log 2 &> H(\phi) \\ \iff H(\psi) + \log 2 + \log d &> H(\phi) + \log d \\ \iff \log d - H(\psi) &< \log 2d - H(\phi) \\ \iff T(\psi) &< T(\phi) \end{aligned}$$

であるから、得られた符号化効率の比較の方法は、 T を用いた符号化効率の比較の方法と一致していることがわかる。このことは、 T の定義の妥当性を示していると言える。

2.3 von Neumann エントロピー

前節で導入した T を使い、KL 変換の次元数と符号化効率の関係について議論するためには、次元数の変化による寄与率分布 ψ の Shannon エントロピー $H(\psi)$ の変化を解析することが必要となる。 $H(\psi)$ の変化は入力信号に依存するものと考えられる。しかし、寄与率分布 ψ が保持するのは、各基底成分の変換係数へのエネルギーの集中度に関する情報のみであり、入力信号に関する情報は保持していない。従って、 ψ のみに着目する場合、 $H(\psi)$ の変化の解析は困難となる。この問題を解決するための数学的道具として、本稿では、以下に示す密度作用素と von Neumann エントロピーを用いる。なお、これらの詳細な説明については、量子情報理論に関する文献（例えば、[8]）を参照されたい。

定義 1 (密度作用素と von Neumann エントロピー) \mathbb{R}^d 上の作用素 ρ が半正定値かつ $\text{tr}\rho = 1$ を満たすとき、 ρ を密度作用素と呼ぶ。密度作用素の固有値 p_1, \dots, p_d に対して、 $p_1, \dots, p_d \geq 0, \sum_{k=1}^d p_k = 1$ が成り立つ。密度作用素 ρ に対して、

$$S(\rho) = -\text{tr}\rho \log \rho = -\sum_{k=1}^d p_k \log p_k$$

を ρ の von Neumann エントロピーと呼ぶ。

定義 2 (部分トレース) ρ をテンソル積空間 $\mathbb{R}^a \otimes \mathbb{R}^b$ 上の密度作用素とし、 $\{e_1, \dots, e_b\}$ を \mathbb{R}^b の正規直交基底とする。このとき、 \mathbb{R}^a 上の密度作用素 σ で、任意の $\xi, \eta \in \mathbb{R}^a$ に対して、 $\langle \sigma \xi, \eta \rangle = \sum_{k=1}^b \langle \rho(\xi \otimes e_k), \eta \otimes e_k \rangle$ を満たすものが存在する。これを $\sigma = \text{tr}_{\mathbb{R}^b} \rho$ と書いて、 ρ の \mathbb{R}^a 上の部分トレースと呼ぶ。特に、

$$\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_b \end{pmatrix} \mapsto \sum_{k=1}^b \xi_k \otimes e_k \quad (\xi_1, \dots, \xi_b \in \mathbb{R}^a)$$

によって与えられる \mathbb{R}^{ab} と $\mathbb{R}^a \otimes \mathbb{R}^b$ の間の同型対応を用いて ρ を \mathbb{R}^{ab} 上の作用素とみなし、 \mathbb{R}^a 上の作用素を要素とする $b \times b$ 行列で表示したものが

$$\begin{pmatrix} A_{11} & \cdots & A_{1b} \\ \vdots & \ddots & \vdots \\ A_{b1} & \cdots & A_{bb} \end{pmatrix}$$

であるとき、部分トレース $\text{tr}_{\mathbb{R}^b} \rho$ は

$$\text{tr}_{\mathbb{R}^b} \rho = \sum_{k=1}^b A_{kk}$$

によって計算することができる。

密度作用素は通常確率論における確率分布の量子力学的な拡張にあたるものであり、von Neumann エントロピーは確率分布に関する Shannon エントロピーの量子力学的な拡張にあたるものである。また、部分トレースは複合系から部分系を取り出す操作と考えることができる。例えば、定義の例では、 $\mathbb{R}^a \otimes \mathbb{R}^b$ により表現される量子確率系を \mathbb{R}^a により表現される部分系と \mathbb{R}^b により表現される部分系との複合系とみなしており、複合系の密度作用素 ρ を \mathbb{R}^a に対応する部分系に制限したものが $\text{tr}_{\mathbb{R}^b} \rho$ となる。

標準基底に関する行列表示によって \mathbb{R}^d 上の作用素を $d \times d$ 行列と同一視すると、密度作用素は半正定値行列 ρ で $\text{tr}\rho = 1$ を満たすものに対応し、その von Neumann エントロピー $S(\rho)$ は ρ の固有値 p_1, \dots, p_d によって得られる確率分布 $\psi = (p_1, \dots, p_d)$ の Shannon エントロピー $H(\psi)$ と一致す

る。このことから、分散共分散行列 V をトレースで正規化することにより得られる密度作用素 $\rho = V/\text{tr}V$ を量子力学的な確率系とみなすことができ、それにより寄与率分布 ψ の Shannon エントロピー $H(\psi)$ についての議論を、密度作用素 ρ の von Neumann エントロピー $S(\rho)$ についての議論に帰着することができる。先に述べた通り、寄与率分布 ψ は変換係数のエネルギーの集中度に関する情報しか持っていない。しかし、密度作用素 ρ は寄与率に加えて入力信号の情報を固有ベクトルという形で含んでいる。これを利用して、画像全体の量子確率系を部分領域 (セグメントと呼ぶ) 単位の量子確率系とセグメント内の画素単位の量子確率系との複合系とみなすと、セグメントを併合する操作を部分トレースにより表現することができる。この部分トレースを用いた表現により、セグメントの併合に伴う von Neumann エントロピー $S(\rho)$ の変化について、解析的な議論が可能となる。

3 理論モデルの導出

本節では 2.2 節で導入した情報集約度 T と KL 変換の次元数との関係を数理的にモデル化する。さらに、同モデルを用いて、入力信号の空間相関が KL 変換の符号化効率に及ぼす影響について考察する。なお、本節では、簡単のため 1 次元信号を対象として議論を行う。

3.1 準備

モデルの導出に先立ち、以降の節で用いる記号について整理する。2dn 個の要素からなるベクトルを x とおき、 x 内の i 番目の要素値を $x(i)$ とする。 d 個の要素 $x(i)$ ($i = dj, \dots, dj+d-1$) から構成される x 内の第 j 番目のセグメントを x_j^d として表す。以降の節では、 x に対する d 次元 KL 変換と $2d$ 次元 KL 変換の関係について考察する。 x に d 次元 KL 変換を施した場合、KL 変換への入力は $2n$ 個のセグメント x_j^d ($j = 0, \dots, 2n-1$) となる。一方、 x に $2d$ 次元 KL 変換を施した場合、KL 変換への入力は n 個のセグメント x_j^{2d} ($j = 0, \dots, n-1$) となる。

さらに、行列形式 X および \tilde{X} を定義する。 $\mathbb{M}_{2n,d}$ を $2n \times d$ 行列の集合として、 $X \in \mathbb{M}_{2n,d}$ を各行が d 次元ベクトル $\sqrt{2}x_j^d$ ($j = 0, \dots, 2n-1$) から構成される $2n \times d$ 行列とする。ここで、 $\sqrt{2}$ は正規化パラメータである。また、 $\tilde{X} \in \mathbb{M}_{n,2d}$ を各行が $2d$ 次元ベクトル x_j^{2d} ($j = 0, \dots, n-1$) から構成される $n \times 2d$ 行列とする。 $X_o \in \mathbb{M}_{n,d}$ を各行が d 次元ベクトル x_{2i-1}^d ($i = 1, \dots, n$) から構成される $n \times d$ 行列とし、 $X_e \in \mathbb{M}_{n,d}$ を各行が d 次元ベクトル x_{2i}^d ($i = 0, \dots, n-1$) から構成される $n \times d$ 行列とする。 X_o と X_e を用いて、 X は次のように表

せる。

$$X = \sqrt{2}\Omega_{2n} \begin{pmatrix} X_e \\ X_o \end{pmatrix}$$

ここで、 Ω_{2n} は、次のように定義される行列である。

$$\begin{cases} i \text{ 列:} & 2i-1 \text{ 行のみ } 1, \text{ それ以外は } 0 \quad (i = 1, \dots, n) \\ j+n \text{ 列:} & 2j \text{ 行のみ } 1, \text{ それ以外は } 0 \quad (j = 1, \dots, n) \end{cases}$$

ここで、 $\begin{pmatrix} X_e \\ X_o \end{pmatrix}$ は、 X の偶数行を上段に、同奇数行を下段に配置した行列である。 Ω_{2n} は、 $\begin{pmatrix} X_e \\ X_o \end{pmatrix}$ に対して行の置換を行い、 X に戻す置換行列である。

3.2 分散共分散行列の導出

d 次元 KL 変換および $2d$ 次元 KL 変換の分散共分散行列を求め、両者の関係を議論する。2.1 節での議論から d 次元 KL 変換の分散共分散行列は、次のように表現できる。

$$\begin{aligned} V &= \frac{1}{2n} X^t E_{2n}^\perp X \\ &= \frac{1}{4n} (X_e^t X_o^t) \Omega_{2n}^t E_{2n}^\perp \Omega_{2n} \begin{pmatrix} X_e \\ X_o \end{pmatrix} \end{aligned} \quad (1)$$

ここで、 Ω_{2n}^t が Ω_{2n} の逆行列であること、さらに、 Ω_{2n} の各行は1つの要素のみ1を含み、それ以外は0であることから、 $\Omega_{2n}^t E_{2n}^\perp \Omega_{2n}$ は次のように展開できる。

$$\begin{aligned} \Omega_{2n}^t E_{2n}^\perp \Omega_{2n} &= \Omega_{2n}^t (I_{2n} - E_{2n}) \Omega_{2n} \\ &= \Omega_{2n}^t I_{2n} \Omega_{2n} - \Omega_{2n}^t E_{2n} \Omega_{2n} \\ &= I_{2n} - E_{2n} \\ &= E_{2n}^\perp \end{aligned} \quad (2)$$

さらに、 E_n が全ての要素を $\frac{1}{n}$ とする $n \times n$ 行列であることから、次式が成り立つ。

$$E_{2n} = \frac{1}{2} \begin{pmatrix} E_n & E_n \\ E_n & E_n \end{pmatrix}$$

あわせて、 $I_n = E_n + E_n^\perp$ であるから、次の関係式を得る。

$$\begin{aligned} E_{2n}^\perp &= I_{2n} - E_{2n} \\ &= \begin{pmatrix} I_n & \mathbf{0} \\ \mathbf{0} & I_n \end{pmatrix} - \frac{1}{2} \begin{pmatrix} E_n & E_n \\ E_n & E_n \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 2I_n - E_n & -E_n \\ -E_n & 2I_n - E_n \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} E_n + 2E_n^\perp & -E_n \\ -E_n & E_n + 2E_n^\perp \end{pmatrix} \end{aligned} \quad (3)$$

式(1)に式(2)および式(3)を代入して、次式が導かれる。

$$\begin{aligned} V &= \frac{1}{4n} (X_e^t X_o^t) E_{2n}^\perp \begin{pmatrix} X_e \\ X_o \end{pmatrix} \\ &= \frac{1}{4n} (X_e^t X_o^t) \begin{pmatrix} E_n + 2E_n^\perp & -E_n \\ -E_n & E_n + 2E_n^\perp \end{pmatrix} \begin{pmatrix} X_e \\ X_o \end{pmatrix} \\ &= \frac{1}{4n} \{ X_e^t (E_n + 2E_n^\perp) X_e - X_o^t E_n X_e \\ &\quad - X_e^t E_n X_o + X_o^t (E_n + 2E_n^\perp) X_o \} \\ &= \frac{1}{4n} \{ 2X_e^t E_n^\perp X_e + 2X_o^t E_n^\perp X_o + X_e^t E_n X_e \\ &\quad - X_o^t E_n X_e - X_e^t E_n X_o + X_o^t E_n X_o \} \\ &= \frac{1}{2n} (X_e^t E_n^\perp X_e + X_o^t E_n^\perp X_o) \\ &\quad + \frac{1}{4n} (X_e - X_o)^t E_n (X_e - X_o) \end{aligned}$$

一方、 $2d$ 次元 KL 変換の分散共分散行列は、以下となる。

$$\begin{aligned} \tilde{V} &= \frac{1}{n} \tilde{X}^t E_n^\perp \tilde{X} \\ &= \frac{1}{n} (X_e X_o)^t E_n^\perp (X_e X_o) \\ &= \frac{1}{n} \begin{pmatrix} X_e^t \\ X_o^t \end{pmatrix} E_n^\perp \begin{pmatrix} X_e \\ X_o \end{pmatrix} \\ &= \frac{1}{n} \begin{pmatrix} X_e^t E_n^\perp X_e & X_e^t E_n^\perp X_o \\ X_o^t E_n^\perp X_e & X_o^t E_n^\perp X_o \end{pmatrix} \end{aligned}$$

このとき、 V と \tilde{V} のトレースは以下となる。

$$\begin{aligned} \text{tr} V &= \frac{1}{2n} \text{tr} (X_e^t E_n^\perp X_e + X_o^t E_n^\perp X_o) \\ &\quad + \frac{1}{4n} \text{tr} \{ (X_e - X_o)^t E_n (X_e - X_o) \} \end{aligned} \quad (4)$$

$$\text{tr} \tilde{V} = \frac{1}{n} \text{tr} \{ (X_e - X_o)^t E_n (X_e - X_o) \} \quad (5)$$

上記の2つの式から、 $\text{tr} V$ と $\text{tr} \tilde{V}$ の関係として、次式を得る。

$$\text{tr} \{ (X_e - X_o)^t E_n (X_e - X_o) \} = 4n \text{tr} V - 2n \text{tr} \tilde{V} \quad (6)$$

3.3 部分トレースを用いた解析

$\rho = \frac{1}{nV} V$ および $\tilde{\rho} = \frac{1}{n\tilde{V}} \tilde{V}$ とおくと、 ρ および $\tilde{\rho}$ は、密度作用素となり、各々、次式のように表せる。

$$\begin{aligned} \rho &= \frac{1}{\text{tr} V} V \\ &= \frac{1}{2n \text{tr} V} (X_e^t E_n^\perp X_e + X_o^t E_n^\perp X_o) \\ &\quad + \frac{1}{4n \text{tr} V} (X_e - X_o)^t E_n (X_e - X_o) \end{aligned} \quad (7)$$

$$\begin{aligned} \tilde{\rho} &= \frac{1}{\text{tr} \tilde{V}} \tilde{V} \\ &= \frac{1}{n \text{tr} \tilde{V}} \begin{pmatrix} X_e^t E_n^\perp X_e & X_e^t E_n^\perp X_o \\ X_o^t E_n^\perp X_e & X_o^t E_n^\perp X_o \end{pmatrix} \end{aligned} \quad (8)$$

このとき、

$$\begin{pmatrix} a_1 \\ \vdots \\ a_d \\ a_{d+1} \\ \vdots \\ a_{2d} \end{pmatrix} \mapsto \begin{pmatrix} a_1 \\ \vdots \\ a_d \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} a_{d+1} \\ \vdots \\ a_{2d} \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (9)$$

によって与えられる \mathbb{R}^{2d} と $\mathbb{R}^d \otimes \mathbb{R}^2$ の間の同型対応を用いて $\tilde{\rho}$ を $\mathbb{R}^d \otimes \mathbb{R}^2$ 上の作用素とみなすことができる。そこで、 $\mathbb{R}^d \otimes \mathbb{R}^2$ 上の混合状態における一部を観測するための数学的な道具として、密度作用素の部分トレースを導入する。 \mathbb{R}^2 上の部分トレースを計算すると、式(8)より

$$\text{tr}_{\mathbb{R}^2} \tilde{\rho} = \frac{1}{n \text{tr} \tilde{V}} (X_e^t E_n^t X_e + X_o^t E_n^t X_o) \quad (10)$$

が導かれる。式(7),(10),(12)を合わせると、

$$\rho = \frac{\text{tr} \tilde{V}}{2 \text{tr} V} \text{tr}_{\mathbb{R}^2} \tilde{\rho} + \frac{2 \text{tr} V - \text{tr} \tilde{V}}{2 \text{tr} V} \sigma \quad (11)$$

が導かれる。ここで、 σ は、 $(X_e - X_o)^t E_n (X_e - X_o)$ をトレースで正規化した作用素であり、次式で表される。

$$\sigma = \frac{1}{4n \text{tr} V - 2n \text{tr} \tilde{V}} (X_e - X_o)^t E_n (X_e - X_o) \quad (12)$$

さらに

$$\frac{\text{tr} \tilde{V}}{2 \text{tr} V} \geq 0, \quad \frac{2 \text{tr} V - \text{tr} \tilde{V}}{2 \text{tr} V} \geq 0, \quad \frac{\text{tr} \tilde{V}}{2 \text{tr} V} + \frac{2 \text{tr} V - \text{tr} \tilde{V}}{2 \text{tr} V} = 1$$

であることから、von Neumann エントロピーの凸性¹より、次式を得る。

$$\begin{aligned} S(\rho) &= S\left(\frac{\text{tr} \tilde{V}}{2 \text{tr} V} \text{tr}_{\mathbb{R}^2} \tilde{\rho} + \frac{2 \text{tr} V - \text{tr} \tilde{V}}{2 \text{tr} V} \sigma\right) \\ &\geq \frac{\text{tr} \tilde{V}}{2 \text{tr} V} S(\text{tr}_{\mathbb{R}^2} \tilde{\rho}) + \frac{2 \text{tr} V - \text{tr} \tilde{V}}{2 \text{tr} V} S(\sigma) \end{aligned} \quad (13)$$

ここで、 $S(\sigma)$ について考察する。 E_n は1次元部分空間への射影であるから、 $(X_e - X_o)^t E_n (X_e - X_o)$ はランクが1の正定値作用素である。従って、 $(X_e - X_o)^t E_n (X_e - X_o)$ をトレースで正規化した σ は、ランクが1の密度作用素となる。 σ のランクが1であることから、von Neumann エントロピーの性質²より $S(\sigma) = 0$ となる。この関係を式(13)に代入し、 $S(\rho)$ は次式となる。

$$S(\rho) \geq \frac{\text{tr} \tilde{V}}{2 \text{tr} V} S(\text{tr}_{\mathbb{R}^2} \tilde{\rho}) \quad (14)$$

¹ ρ_1, ρ_2 は密度作用素で、 $a_1, a_2 \geq 0$, $a_1 + a_2 = 1$ であるとき、次の関係が成り立つ。

$$S(a_1 \rho_1 + a_2 \rho_2) \geq a_1 S(\rho_1) + a_2 S(\rho_2)$$

² \mathbb{R}^d 上の密度作用素 ρ の von Neumann エントロピーに対して、 $0 \leq S(\rho) \leq \log d$ が成り立つ。 $S(\rho) = 0$ となるのは ρ のランクが1であるときに限る。

さらに von Neumann エントロピーの劣加法性³から

$$S(\tilde{\rho}) \leq S(\text{tr}_{\mathbb{R}^d} \tilde{\rho}) + S(\text{tr}_{\mathbb{R}^2} \tilde{\rho})$$

が成り立つので、式(14)と合わせることで、

$$\begin{aligned} S(\tilde{\rho}) &\leq S(\text{tr}_{\mathbb{R}^d} \tilde{\rho}) + \frac{2 \text{tr} V}{\text{tr} \tilde{V}} S(\rho) \\ &= S(\rho) + S(\text{tr}_{\mathbb{R}^d} \tilde{\rho}) + \frac{2 \text{tr} V - \text{tr} \tilde{V}}{\text{tr} \tilde{V}} S(\rho) \end{aligned} \quad (15)$$

が導かれる。

上式の右辺の第三項について考察する。行列 X_e, X_o の第 k 列をそれぞれ $x_e^{(k)}, x_o^{(k)}$ と表すことにする。すなわち

$$X_e = (x_e^{(1)} \cdots x_e^{(d)}), \quad X_o = (x_o^{(1)} \cdots x_o^{(d)})$$

とする。このとき、 $E_n(X_e - X_o) = (E_n(x_e^{(1)} - x_o^{(1)}) \cdots E_n(x_e^{(d)} - x_o^{(d)}))$ と表せること、および、式(6)の関係を用いると、式(15)の第三項は次のように展開できる。

$$\begin{aligned} 2 \text{tr} V - \text{tr} \tilde{V} &= \frac{1}{2n} \text{tr} \left\{ (X_e - X_o)^t E_n (X_e - X_o) \right\} \\ &= \frac{1}{2n} \sum_{k=1}^d \|E_n x_e^{(k)} - E_n x_o^{(k)}\|^2 \end{aligned} \quad (16)$$

ここで、入力信号に次の性質を仮定する。

性質1 X_o に含まれるセグメント x_{2i-1}^d ($i = 0, \dots, n$) および X_e に含まれるセグメント x_{2i}^d ($i = 0, \dots, n$) 内の同じ相対位置にある画素の平均値が、その相対位置によらず一定である。

性質1は、次のように言い換えることができる。

$$E_n x_e^{(k)} = E_n x_o^{(k)} = C \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \quad (C \text{ は定数。} k = 1, \dots, d)$$

本論文では、KL 変換の入力信号としてフレーム間予測誤差信号を用いることを想定している。こうした予測誤差信号の平均値は零値であることが期待できるため、上記の性質を仮定することは無理の無いものとする。上記の性質を用いて、次式を得る。

$$\sum_{k=1}^d \|E_n x_e^{(k)} - E_n x_o^{(k)}\|^2 = 0$$

上式および式(16)より、次式を得る。

$$2 \text{tr} V - \text{tr} \tilde{V} = 0$$

従って、性質1を仮定するとき、式(15)の右辺第三項は0となることがわかり、次式が導かれる。

$$S(\tilde{\rho}) \leq S(\rho) + S(\text{tr}_{\mathbb{R}^d} \tilde{\rho}) \quad (17)$$

³ $\mathbb{R}^a \otimes \mathbb{R}^b$ 上の密度作用素 ρ に対して、次の関係が成り立つ。

$$S(\rho) \leq S(\text{tr}_a \rho) + S(\text{tr}_b \rho)$$

3.4 情報集約度を用いた評価

d 次元 KL 変換および $2d$ 次元 KL 変換の寄与率分布を各々 $\psi, \tilde{\psi}$ とする。

$$T(\psi) = \log d - H(\psi) = \log d - S(\rho),$$

$$T(\tilde{\psi}) = \log 2d - H(\tilde{\psi}) = \log 2d - S(\tilde{\rho})$$

であるから、式 (17) と合わせることにより、2種類の KL 変換の情報集約度に関して次の不等式が得られる。

$$T(\psi) + \log 2 - S(\text{tr}_{\mathbb{R}^d} \tilde{\rho}) \leq T(\tilde{\psi}) \quad (18)$$

左辺において $\log 2 - S(\text{tr}_{\mathbb{R}^d} \tilde{\rho}) \geq 0$ が成り立つことから、式 (18) は、「入力信号が条件 1 を満たすならば KL 変換の次元数を 2 倍にすることによって KL 変換の符号化効率率は必ず向上する」ことを示していると解釈できる。

入力信号の空間的相関の観点からさらに考察を行うため、入力信号の空間的相関と $\text{tr}_{\mathbb{R}^d} \tilde{\rho}$ の関係を考察する。紙面の都合上、詳細は割愛するが、 $\text{tr}_{\mathbb{R}^d} \tilde{\rho}$ は、次のように展開できる。

$$\begin{aligned} \text{tr}_{\mathbb{R}^d} \tilde{\rho} &= \frac{1}{n\text{tr}\tilde{V}} \text{tr}_{\mathbb{R}^d} \begin{pmatrix} \mathbf{X}_e^t \mathbf{E}_n^+ \mathbf{X}_e & \mathbf{X}_e^t \mathbf{E}_n^+ \mathbf{X}_o \\ \mathbf{X}_o^t \mathbf{E}_n^+ \mathbf{X}_e & \mathbf{X}_o^t \mathbf{E}_n^+ \mathbf{X}_o \end{pmatrix} \\ &= \frac{1}{n\text{tr}\tilde{V}} \sum_{k=1}^d \begin{pmatrix} \mathbf{x}_e^{(k)t} \mathbf{E}_n^+ \mathbf{x}_e^{(k)} & \mathbf{x}_e^{(k)t} \mathbf{E}_n^+ \mathbf{x}_o^{(k)} \\ \mathbf{x}_o^{(k)t} \mathbf{E}_n^+ \mathbf{x}_e^{(k)} & \mathbf{x}_o^{(k)t} \mathbf{E}_n^+ \mathbf{x}_o^{(k)} \end{pmatrix} \end{aligned}$$

このとき、 $r_{oo}(d) = \frac{1}{n\text{tr}\tilde{V}} \sum_{k=1}^d \mathbf{x}_o^{(k)t} \mathbf{E}_n^+ \mathbf{x}_o^{(k)}$ 、 $r_{oe}(d) = \frac{1}{r_{oo}(d)n\text{tr}\tilde{V}} \sum_{k=1}^d \mathbf{x}_e^{(k)t} \mathbf{E}_n^+ \mathbf{x}_o^{(k)}$ 、 $r_{ee}(d) = \frac{1}{n\text{tr}\tilde{V}} \sum_{k=1}^d \mathbf{x}_e^{(k)t} \mathbf{E}_n^+ \mathbf{x}_e^{(k)}$ とおくと、 $\text{tr}_{\mathbb{R}^d} \tilde{\rho}$ は次のように表せる。

$$\text{tr}_{\mathbb{R}^d} \tilde{\rho} = \begin{pmatrix} r_{ee}(d) & r_{ee}(d)r_{eo}(d) \\ r_{ee}(d)r_{oe}(d) & r_{oo}(d) \end{pmatrix}$$

ここで、近似式 $r_{oo}(d) \approx r_{ee}(d) \approx \frac{1}{2}$ を導入する。さらに、この近似に加えて、 $r_{oe}(d) = r_{eo}(d)$ および $\text{tr}\tilde{V} = \frac{1}{n} \{ \sum_{k=1}^d \mathbf{x}_e^{(k)t} \mathbf{E}_n^+ \mathbf{x}_e^{(k)} + \sum_{k=1}^d \mathbf{x}_o^{(k)t} \mathbf{E}_n^+ \mathbf{x}_o^{(k)} \}$ に注意すると、 $\text{tr}_{\mathbb{R}^d} \tilde{\rho}$ の固有値は $\frac{1 \pm |r_{ee}(d)|}{2}$ となる。 $\gamma_d = (\frac{1 - |r_{ee}(d)|}{2}, \frac{1 + |r_{ee}(d)|}{2})$ とおくと、 $\text{tr}_{\mathbb{R}^d} \tilde{\rho}$ の von Neumann エントロピーは、 $S(\text{tr}_{\mathbb{R}^d} \tilde{\rho}) = H(\gamma_d)$ となる。従って $\log 2 - S(\text{tr}_{\mathbb{R}^d} \tilde{\rho})$ は一様分布 u_2 に対する確率分布 $\gamma_d = (\frac{1 - |r_{ee}(d)|}{2}, \frac{1 + |r_{ee}(d)|}{2})$ の Kullback-Leibler divergence、即ち、情報集約度に等しいことがわかる。そこで、 $\tau(d) = \log 2 - S(\text{tr}_{\mathbb{R}^d} \tilde{\rho})$ とおき、式 (18) を次のように表す。

$$T(\psi) + \tau(d) \leq T(\tilde{\psi}) \quad (19)$$

以下では、 $\tau(d)$ と d の関係について考察する。まず、 $r_{eo}(d)$ と d の関係について整理する。 $r_{eo}(d)$ は $\mathbf{x}_o^{(k)}$ と $\mathbf{x}_e^{(k)}$ との間のサンプル相互相関であり、 $\mathbf{x}_o^{(k)}$ および $\mathbf{x}_e^{(k)}$ の第 j 要素は、各々、 $\mathbf{x}(d(2j-1)+k)$ および $\mathbf{x}(2dj+k)$ となることに注意す

表 1: 動き推定 (ME) のパラメータ

Items	Parameters
ME algorithm	block matching base
ME block size	16 × 16
ME search range	± 32 × 32
ME accuracy	integer pixel
Total number of references	1
ME scheme	Full search
ME Cost	Sum of square error

ると、 $r_{eo}(d)$ は入力信号 \mathbf{x} に対するサンプル自己相関 (シフト量 d) であることが分かる。

一般に、画像信号の自己相関はシフト量に対して減少関数であるため、 $r_{eo}(d)$ は d に対する減少関数とみなすことができる。このため、 d が増加すれば、 γ_d は $(\frac{1}{2}, \frac{1}{2})$ に漸近し、結果として、 $H(\gamma_d)$ は $\log 2$ に漸近する。つまり、KL 変換の次元数 d が増加する場合、 $\tau(d)$ は減少することになる。式 (19) より、 $\tau(d)$ は d 次元 KL 変換と $2d$ 次元 KL 変換の情報集約度の差分値の下限值とみなせる。従って、上述の $\tau(d)$ と d の関係から、式 (19) は、「KL 変換の次元数が増加するにつれ、KL 変換の情報集約度が増加する割合は鈍化する」ことを示していると解釈できる。

また、 $r_{eo}(d)$ が入力信号のサンプル自己相関であることから、入力信号の空間的相関が高いほど $r_{eo}(d)$ は大きな値をとり、結果として、 $\tau(d)$ が大きな値をとることが予想される。このため、式 (19) は、「KL 変換の次元数を増加させた際、KL 変換の情報集約度が増加する割合は、画素間の相関が高い入力信号ほど高い」ことを示しているといえる。

4 実験

前節で導出した理論モデルの妥当性を検証するために以下の実験を行った。本実験では、KL 変換の入力信号としてフレーム間予測誤差信号を用いた。フレーム間予測はブロックマッチングベースの動き補償に基づく方法である。同フレーム間予測で用いたパラメータを表 1 にまとめる。フレーム間予測の対象となる動画シーケンス (City, Crew) は、352 × 288 [画素/フレーム]、30 [フレーム/秒] のものを用いた。

図 1 に、KL 変換の再構成に用いた基底ベクトルと累積寄与率の関係を示す。KL 変換の次元数は、4 × 4、8 × 8、16 × 16 の 3 種類とした。同図の縦軸は、基底ベクトルの総数 (KL 変換の次元) に対する再構成に用いられた基底ベクトルの数の比を示している。横軸は、再構成に用いられた基底ベクトルの固有値の寄与率を表す。同図は、累積寄与率の観

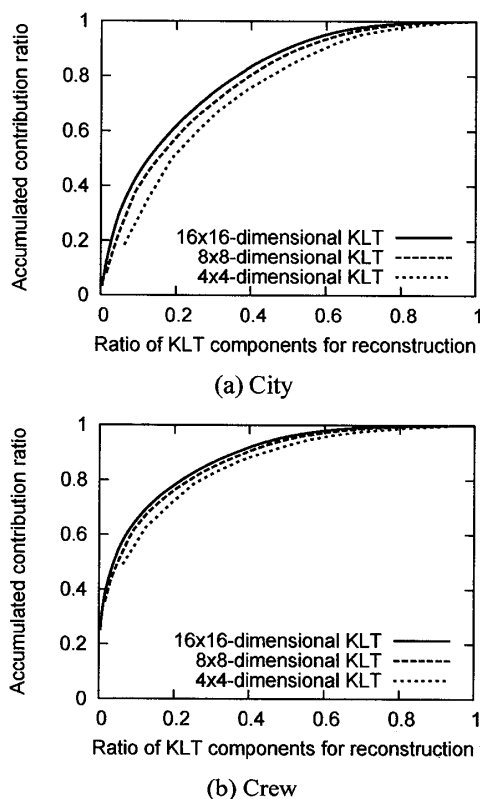


図1: KL変換の次元数と累積寄与率分布

点から、KL変換の情報集約性能を示したものである。あわせて、表2に、3種類のKL変換（次元数 4×4 、 8×8 、 16×16 ）に対するKL変換の情報集約度の値を示す。

図1より、KL変換の次元数が増加するとグラフが左上に移動していることが確認できる。これは、KL変換により得られる情報集約度が増加していることを意味する。表2からも、KL変換の次元数の増加に対して、情報集約度が増加傾向にあることを確認できる。これらの観測結果は、KL変換の次元数が増加するにつれて、符号化効率という提案モデルの性質と一致する。

さらに、図1において、グラフの左上への移動量は、KL変換の次元数の増加に伴い、減少していくことが見て取れる。これは、情報集約度の増加率が小さくなっていくことを意味する。表2からも、KL変換の次元数の増加に対して、情報集約度の増加率が鈍化傾向にあることを確認できる。これらの観測結果は、KL変換の次元数が増加するにつれて、KL変換の符号化効率が向上する効果は鈍化するという提案モデルの性質と一致する。

5 おわりに

本稿では、KL変換の次元数とその情報集約度に関する数理モデルを導出し、両者の関係を明らかにした。導出し

表2: 情報集約度

	4×4次元 KL変換	8×8次元 KL変換	16×16次元 KL変換
City	0.34	0.47	0.58
Crew	0.79	1.06	1.27

たモデルにより、KL変換の次元数を増加させた場合、KL変換の情報集約度が増加し、かつ、その増加率が収束傾向にあることを理論的に示すことができた。また、実画像を用いた検証実験において、上述のモデルの性質と合致する実験結果を観測し、導出したモデルの妥当性を確認することができた。

参考文献

- [1] G. Sullivan, P. Topiwala and A. Luthra: "The H.264/AVC advanced video coding standard: Overview and introduction to the fidelity range extensions", Proc. SPIE Applications of Digital Image Processing XXVII, Vol. 5558, pp. 454-474 (2004).
- [2] T. Yamamoto and T. Ikai: "A new intra macroblock type for adaptable transform block size", ITU-T Q.6/SG16 Doc., COM 16-C 88-E., (2009).
- [3] A. K. Soman and P. P. Vaidyanathan: "Coding gain in paraunitary analysis/synthesis systems", IEEE Trans. Signal Processing, **41**, 5, pp. 1824 - 1835 (1993).
- [4] N. Jayant and P. Noll: "Digital Coding of Waveforms", Prentice Hall (1984).
- [5] J. P. Andrew, P. O. Ogunbona and F. J. Paoloni: "coding gain and spatial localisation properties of discrete wavelet-transform filters for image coding", IEE Proc. Vision, Image and Signal Processing, pp. 133-140 (1995).
- [6] P. Vaidyanathan: "Multirate Systems and Filter Banks", Prentice Hall (1992).
- [7] T. Cover and J. Thomas: "Elements of Information Theory", Wiley-Interscience (2006).
- [8] M. A. Nielsen and I. L. Chuang: "Quantum Computation and Quantum Information", Cambridge Univ. Press (2000).