

# Une application comparative pour l'enseignement des méthodes de compression d'images par transformées orthogonales

Mohamed Gharbi<sup>1</sup>, Mohamed Aymen Labiod<sup>1</sup>

Mohamed.Gharbi@uphf.fr

(1) Université Polytechnique Hauts de France, IEMN UMR CNRS 8520

**RESUME** : Nous proposons dans cette contribution une interface graphique comme outil pédagogique pour l'enseignement des méthodes de compression par transformées orthogonales utilisées dans l'encodage d'images. Cet outil est utilisé dans l'enseignement proposé aux étudiants en MASTER du département Développement, Recherche, Enseignement en Audiovisuel et Multimédia (DREAM) de l'Université Polytechnique Hauts-de-France (Valenciennes – France). Les offres de formation s'appuient notamment sur un enseignement du traitement de signal, des images ainsi que la compression des images et de la vidéo. Les deux aspects analogique et numérique du traitement du signal sont enseignés. Cet interface graphique a été développé pour permettre aux étudiants de mieux appréhender l'influence des différents paramètres intervenant dans une chaîne de compression numérique des images. A travers cet outil, nous proposons aux étudiants une nouvelle approche à la compréhension des transformées matricielles utilisées dans l'encodage des images et de la quantification sur la qualité des images compressées par plusieurs transformées. En effet, cet outil, par le biais de l'ajustement des différents paramètres intervenant dans le chaîne de compression, permet aux étudiants d'étudier et de comparer plusieurs transformées orthogonales à savoir la Transformée en Cosinus Discrète (DCT), la Transformée de Karhunen-Loève (KLT) sous les deux versions séparable ou non, la Transformation de Fourier Discrète (TFD) ainsi que la Transformée d'Hadamard. Ainsi, visuellement par le biais de l'image reconstruite et par le rapport signal à bruit cet outil permet aux étudiants d'apprécier le choix des paramètres qu'ils ont fixés pour l'obtention de leurs résultats.

**Mots clés** : Transformées, encodage d'images, Transformée en Cosinus Discrète (DCT), Transformée de Karhunen-Loève (KLT), Transformation de Fourier rapide (FFT).

## 1 INTRODUCTION

L'Université Polytechnique Hauts-de-France possède une solide réputation dans le domaine de la formation professionnelle aux métiers de l'audiovisuel et des médias numériques [1]. Notamment avec l'offre de Master qui s'appuie sur trois parcours différents au sein du département DREAM (Développement, Recherche, Enseignement en Audiovisuel et Multimédia). Le département prépare les étudiants aux métiers de l'image et du son et cela du Baccalauréat (BAC) jusqu'au doctorat à BAC+8. L'offre de Master propose plusieurs parcours dont, entre autres, le parcours type Trucage Numérique des Images et des Sons (TRUCIS) et le parcours type Ingénierie de Systèmes Images et Sons (ISIS).

Pour le parcours ISIS, plus orienté vers les métiers techniques du multimédia, les méthodes compression des images et de la vidéo utilisées dans beaucoup de secteurs et plus particulièrement par tous les diffuseurs télévisuels. La plupart des standards actuels de compression des images et de la vidéo sont basés sur les méthodes utilisant les méthodes par transformées. Les transformations utilisées font partie de la boîte à outil indispensable du traitement des signaux et des images. Elles utilisent les résultats mathématiques sur les matrices et l'algèbre linéaire. D'un autre côté, l'enseignement des transformées a la réputation d'être difficile parce qu'enraciné dans un formalisme mathématique. En effet, nous constatons que d'une manière plus globale la méthode d'enseignement de la

théorie de l'information et du traitement du signal doit utiliser de plus en plus avec les moyens actuels disponibles de traitement et de simulation. D'autant plus qu'aujourd'hui de nouvelles formes d'enseignement : classe inversée, enseignement à distance, par projet, etc. Dans cet article, nous exposons les deux parties de cet enseignement : la partie classique en classe et celle utilisée par cet interface graphique d'utilisateur (GUI). Cet enseignement comporte dans sa partie classique une introduction théorique didactique menant à appréhender les transformations orthogonales [1,2]. Dans sa partie basée sur la simulation, nous proposons aux étudiants d'étudier et de comparer les méthodes de compression par quantification zonale utilisant les transformées suivantes : en Cosinus Discrète (DCT), de Karhunen-Loève (KLT), de Fourier Discrète (TFD) ainsi que la Transformée d'Hadamard.

Le reste de ce papier est organisé comme suit. La Section 2 présente une brève introduction à la compression des signaux pour ensuite se focaliser dans la section 3 sur la description des différentes transformées orthogonales étudiées. La section 4 concerne les transformées dans le cas de l'encodage d'images avec une présentation des ressources numériques pédagogiques utilisés pour l'apprentissage. Enfin, une conclusion est discutée dans la section 5.

## 2 INTRODUCTION A LA COMPRESSION DES SIGNAUX

On explique aux étudiants le but de la compression des signaux : la réduction de la représentation du signal à

une forme plus compacte et donc, si possible, à la taille (en bits) la plus faible nécessaire au codage de l'information contenue dans le signal. On sait d'après la forme autorégressive du théorème de Wold que tout processus aléatoire peut s'écrire comme somme d'une partie purement prévisible et d'un bruit blanc (innovation) [5, page 27]. Pour réduire la taille de la représentation du signal, il faudra théoriquement ne conserver que l'innovation, c'est à dire la partie non prédictible du signal : il est superflu de stocker la partie prévisible.

Comme les images sont très corrélées (corrélation spatiale entre pixels voisins), il est nécessaire de passer par une étape de décorrélation pour ne garder que le signal d'innovation. Plusieurs techniques sont possibles dont l'utilisation des transformées réversibles qui permettent de compacter la représentation en décorrélant les données. Une des transformations linéaires qui a prouvé son efficacité dans le domaine de la compression des images est la Transformée en Cosinus Discrète (DCT) qui à la fois possède un algorithme rapide de calcul et arrive dans une certaine mesure à compacter l'information.

### 3 LES TRANSFORMATIONS ORTHOGONALES

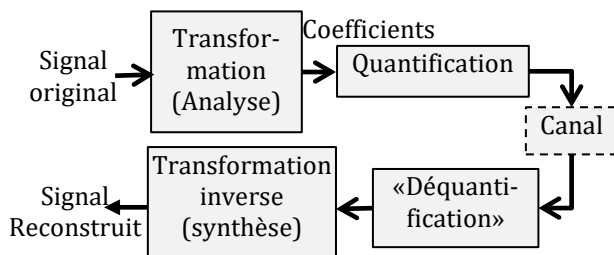


Figure 1. Schéma simplifié du système de compression.

#### 3.1 Transformation de Fourier discrète (DFT)

L'enseignement de la transformée de Fourier a déjà été effectué au niveau de la licence (Bac+3). Il s'agit de la bien entendu de l'étude des séries de Fourier, de la transformée de Fourier des signaux à temps continu et discrets ainsi qu'une introduction à la transformée de Fourier Discrète (TFD) et son algorithme de calcul rapide (FFT). Pour l'enseignement de niveau Master considéré dans cet article seule la TFD est reprise et est reformulée en termes de changement de base et donc de transformation matricielle. Le parallèle avec l'analyse vectorielle est bien entendu utilisé aussi souvent que possible permettant aux étudiants de toujours se ramener à l'univers connu et maîtrisé des vecteurs.

#### 3.2 Les transformations orthogonales 1D

le formalisme matriciel de la transformée de Fourier est généralisé aux transformations orthogonales quelconques. Considérons  $N$  échantillons d'un signal unidimensionnel représenté sous forme du vecteur  $x$  de taille  $N \times 1$ . Soit  $A$  la matrice de transformation, le vecteur  $X$  constitué des coefficients transformés est

obtenu par  $X = A \cdot x$ . Si  $A$  est orthogonale (unitaire), i.e. vérifiant  $A^{-1} = A^H$ , où l'exposant  $H$  (symétrie hermitienne) désigne l'opération de transposition  $T$ , suivie de la conjugaison complexe  $*$ , l'opération est réversible et le signal original peut être retrouvé à partir du vecteur  $X$  par  $x = A^H \cdot X$ .

L'orthogonalité de la matrice  $A$  assure la conservation de l'énergie dans le domaine transformé, c'est-à-dire,  $\|X\|^2 = \|x\|^2$ .

Pour s'adapter à la statistique locale du signal, le signal est souvent partitionné en blocs, généralement de même taille, et la transformation opère bloc par bloc et sans recouvrement des blocs. C'est aussi le cas de l'outil proposé dans cet article qui considère 15 tailles de blocs allant de  $2 \times 2$  à  $16 \times 16$ .

#### 3.3 Les transformations orthogonales 2D

Un signal bidimensionnel est représenté par une matrice, par exemple une image de  $M$  lignes et  $N$  colonnes. Il faut considérer dans le cas des signaux bidimensionnels, deux classes de transformation : les transformations séparables et les non-séparables. La propriété de séparabilité permet de décomposer la transformation 2D en une opération sur les lignes (ou les colonnes) de la matrice représentant l'image suivie d'une transformation sur les colonnes (resp. les lignes). Quand la propriété de séparabilité n'est pas vérifiée, une représentation matricielle peut être obtenue en vectorisant le bloc à traiter. Comme la transformation est linéaire, le signal bidimensionnel peut être vectorisé, c'est à dire mis sous forme d'un vecteur de taille  $M \cdot N \times 1$ . Par exemple, avec le langage Matlab par l'instruction  $x(:)$ . La transformation non-séparable sera représentée par une matrice, de taille  $M \cdot N \times M \cdot N$  qui agissant sur l'image vectorisée produit le signal bidimensionnel transformé vectorisé de taille  $M \cdot N \times 1$ .

#### 3.4 Transformée en Cosinus Discrète (DCT)

Utilisant la propriété de symétrie hermitienne de la transformée de Fourier concernant les signaux réels, les transformées en cosinus ou en sinus discrètes sont introduites. Comme il existe plusieurs manière de symétriser une séquence de taille finie à chacune de ses deux extrémités, plusieurs transformations peuvent alors être engendrées. Pour les symétries impaires on obtient les transformées en sinus discrètes et pour les symétries paires les transformées en cosinus discrètes. Les matrices de transformation associées sont réelles.

La DCT est la transformation la plus utilisée dans la compression des signaux audionumériques, les images et la vidéo. En compression des images c'est le type 2 de la DCT qui est utilisé. A l'instar de la DFT, elle conserve la propriété de réduire à un seul coefficient tout signal constant. Elle dispose, comme la transformée de Fourier d'un algorithme de calcul rapide.

#### 3.5 Transformée de Karhunen-Loève (KLT)

Comme expliqué précédemment, les performances des DCT, sont particulièrement importantes dans le codage

par transformation, et sont souvent associées à la Transformée de Karhunen-Loève (KLT). En effet, la KLT est la transformation statistiquement optimale pour la compression de données car elle génère des coefficients transformés décorrélés, ce qui lui permet de compacter l'information contenue dans le signal dans un nombre réduit de coefficients. Ainsi, on peut montrer que si on approxime le signal original en ne gardant qu'une partie des coefficients l'erreur quadratique moyenne d'approximation se trouve plus faible que celle obtenue par n'importe quelle autre transformation linéaire. Cependant, la KLT est constituée des vecteurs propres de la matrice de covariance des données à transformer. Contrairement aux autres transformations citées, elle est "signal-dépendante" dans le sens où la matrice de transformation KLT dépend du signal sur lequel elle agit. A chaque signal d'entrée une matrice de transformation est associée, il est donc difficile de trouver dans le cas général un algorithme de calcul rapide des coefficients transformés. La KLT est souvent utilisée seulement comme « benchmark » à titre de comparaison.

Il existe une équivalence asymptotique de la famille des DCT avec la KLT pour un processus de Markov stationnaire de premier ordre en termes de taille de transformation et de coefficient de corrélation  $\rho$  (entre échantillons adjacents) [1-3]. Pour les données de longueur finie, les DCT fournissent des approximations différentes de la KLT, et la meilleure transformée d'approximation dépend de la valeur du coefficient de corrélation  $\rho$ .

Par exemple, lorsque  $\rho = 1$ , la KLT se réduit à la DCT-II (DCT-III), pour  $\rho = 0$ , à la DST-I, et pour  $\rho = -1$ , à DST-II (DST-III). En revanche, si la taille de transformation  $N$  augmente (c'est-à-dire que  $N$  tend vers l'infini), on montre que la KLT se réduit à la DCT-I ou DCT-IV. Ce comportement asymptotique implique que

les DCT et les DST peuvent être utilisées comme substituts à la KLT.

Dans la famille des DCT et des DST, les performances de la DCT-II sont plus proches de la KLT en termes de plusieurs critères de performance comme par exemple : distribution de la variance des coefficients, efficacité du compactage énergétique, corrélation résiduelle. L'importance de la DCT-II est encore accentuée par l'existence d'algorithmes rapides pour sa mise en œuvre [3-4]. Pour cela, DCT-II a été utilisée dans pratiquement toutes les normes internationales de codage image/vidéo : JPEG pour la compression d'images fixes, plusieurs normes MPEG de compression vidéo, ainsi que dans les différents codec vidéo du H.261 au H.265/HEVC [6].

### 3.6 Transformée de Hadamard

La transformation de Hadamard (également connue sous le nom de Walsh-Hadamard ou encore de Walsh) est orthogonale, symétrique, involutive (égale à son inverse). Les éléments de la matrice de transformation sont égaux à  $\pm 1$ , la transformation ne requiert donc aucune multiplication (à une constante de normalisation près). La matrice de Hadamard peut être définie de deux manières : récursivement, ou en utilisant la représentation binaire (base 2) des indices  $n$  et  $k$  des indices lignes et colonnes. Un algorithme rapide de calcul existe pour toute taille puissance de 2.

## 4 INTERFACE GRAPHIQUE SUR LES TRANSFORMÉES POUR L'ENCODAGE D'IMAGES (2D)

Opérant sur une image choisie par l'utilisateur, parmi des images conventionnelles ou images pédagogiques à contenu fréquentiels connu et fixé, l'interface graphique réalisé permet aux étudiants de comparer les performances de compression obtenus avec plusieurs matrices de transformation. Les résultats sont

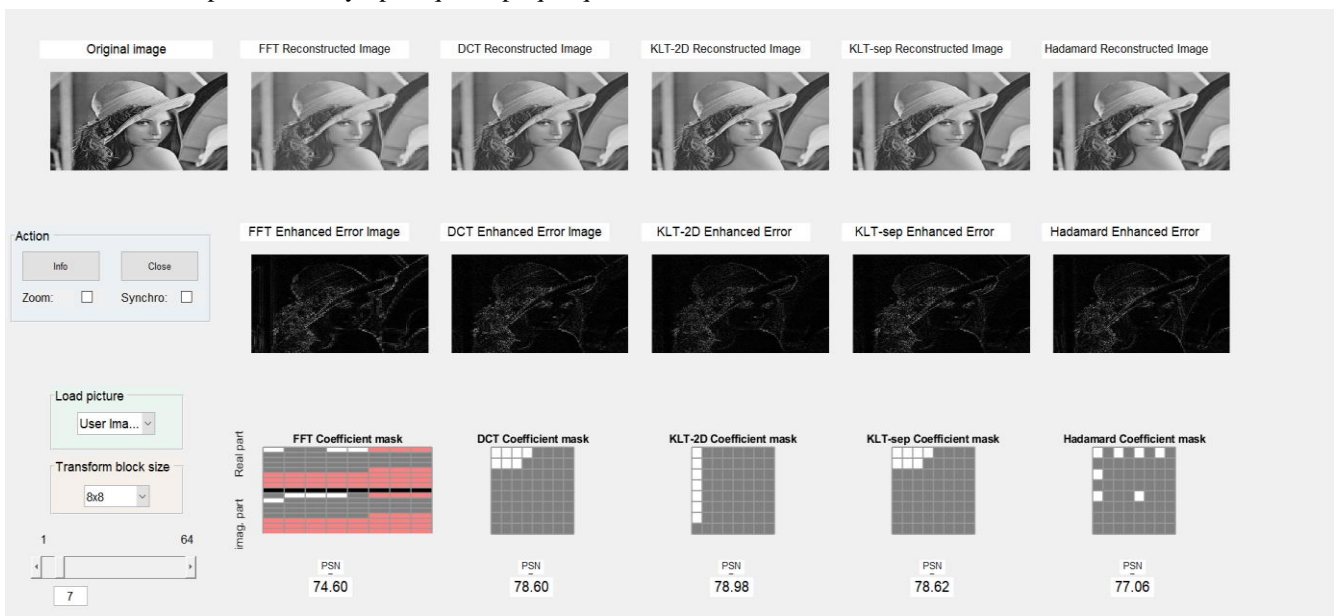


Figure 2. Interface graphique utilisateur de l'application méthodes par transformées.

représentés sur une même fenêtre : l'image transformée, l'image d'erreur et le PSNR. Les transformations utilisées sont celles de : Fourier, la DCT2, la KLT-2D, la KLT-2D séparable et Hadamard. Seule la technique de quantification zonale a été implémentée (zonal sampling quantization technique) Fig. 1. Cette technique opérant bloc par bloc consiste à annuler des coefficients transformés dont l'énergie totale calculée sur toute l'image est la plus faible. Le nombre de coefficients gardés est fixé par l'utilisateur en agissant sur l'ascenseur (slider) visible au bas à gauche de la fenêtre. Ce nombre est compris entre 1 et la taille des blocs utilisés par les transformations. La taille des blocs est choisie par l'utilisateur :  $2 \times 2$ ,  $3 \times 3$ , ...,  $16 \times 16$ . La taille maximale a été délibérément limitée pour que le programme puisse tourner même sur des machines à performances réduites. A noter que la transformée de Hadamard n'est valable que pour une taille puissance de 2.

Sur la partie gauche de l'interface (Fig. 2), les étudiants disposent de l'aide utilisateur (bouton info) et la possibilité de changer d'image, de zoomer indépendamment sur chaque image ou en synchronisme sur toutes les images pour pouvoir comparer finement la qualité d'une partie des différentes images obtenues. Pour que la comparaison de ces différentes méthodes soit valable, un même nombre de coefficients réels est conservé après l'étape de quantification zonale pour toutes les transformations. Un traitement particulier concerne la TFD qui délivre des coefficients complexes. Comme l'image originale est réelle, les coefficients

transformés possèdent la symétrie hermitienne. Pour une taille de bloc  $N \times N$ , on a :

$$X(N-n, N-k) = X^*(n, k), \quad n, k = 0, 1, \dots, N-1.$$

Pour une valeur, par exemple paire de  $N$ , on a donc  $X(0,0)$ ,  $X(0, N/2)$ ,  $X(N/2, 0)$  et  $X(N/2, N/2)$  réels. Pour éviter toute redondance, on ne considère pour les parties réelles,  $\mathcal{R}\{\cdot\}$ , des coefficients :

$$\mathcal{R}\{X(0, k)\}, \quad n = 0, \frac{N}{2}, \quad k = 0, 1, \dots, \frac{N}{2}.$$

$$\mathcal{R}\{X(n, k)\}, \quad n = 1, 2, \dots, \frac{N}{2} - 1, \quad k = 0, 1, \dots, N-1.$$

Soit  $\frac{N^2}{2} + 2$  coefficients. Pour les parties imaginaires,  $\mathcal{I}\{\cdot\}$ , on ne considère que :

$$\mathcal{I}\{X(0, k)\}, \quad n = 0, \frac{N}{2}, \quad k = 1, 1, \dots, \frac{N}{2} - 1.$$

$$\mathcal{I}\{X(n, k)\}, \quad n = 1, 2, \dots, \frac{N}{2} - 1, \quad k = 0, 1, \dots, N-1.$$

Soit  $\frac{N^2}{2} + 2$  coefficients. Au total donc Soit  $N^2$  coefficients réels et donc le nombre de pixels du bloc.

La répartition obtenue des coefficients gardés par chaque transformation est visualisée pour permettre aux étudiants de localiser les coefficients les plus énergiques obtenus par chaque transformation. Comme généralement convenu le coefficient DC (fréquences horizontales et verticales nulles) est situé en haut à gauche du bloc représenté. A titre d'exemple, comme attendu, on remarque le compactage de l'information de la DCT, i.e. les coefficients obtenus sont localisés au haut à gauche du bloc souvent ressemblant à celui de la KLT-séparable. Pour la KLT-2D l'ordre visualisé n'a aucune de signification.

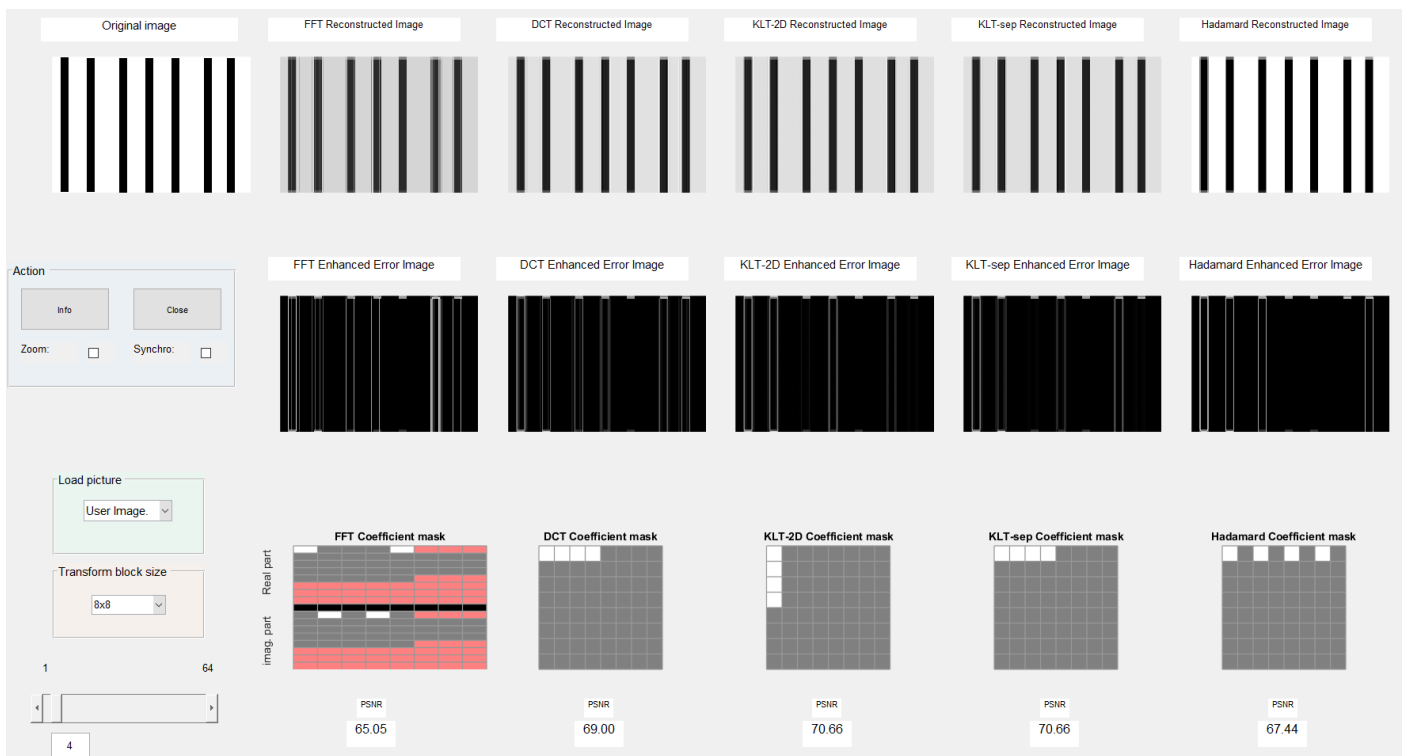


Figure 3. Exemple avec une image composée de barres verticales.

## 5 CONCLUSION

Dans cet article, nous avons présenté une approche de l'enseignement de la compression d'images par transformée qui s'appuie sur un interface graphique permettant de comparer plusieurs transformations orthogonales à savoir la DCT type2, la KLT 2D , la KLT séparable, la transformation de Fourier discrète ainsi que la transformée d'Hadamard. En effet, les étudiants pourront ainsi mieux appréhender l'influence des différents paramètres en intervenant sur l'image elle-même, la taille des blocs ainsi que le nombre de coefficients retenu. Les images transformées sont affichées dans une même fenêtre.

### Bibliographie

- [1] K. R. Rao and P.C. Yip, "The transform and data compression handbook", *Boca Raton, CRC Press 2000*.
- [2] R. C. Gonzalez, R. E. Woods, "Digital Image Processing", *4th Edition , editor Pearson, 2018*.
- [3] K. R. Rao, " Discrete Cosine Transform: Algorithms, Advantages, Applications ", *Academic Press Inc, 2006*.
- [4] G. Plonka, M. Tasche, "Fast and numerically stable algorithms for discrete cosine transforms" . *Linear Algebra and Its Applications Vol. 394, 1 Jan. 2005, Pages 309-345*.
- [5] Boaz Porat, "Digital Processing of Random Signals: Theory and Methods", *Ed. Dover Publications Inc., 2008*.
- [6] W. Fischer, "Digital Video and audio broadcasting technology: A Practical Engineering Guide", *Springer, 2020*.