Chapter 3

Wavelets

- *Focus of thesis is on Images and in particular Videos*

- *Why is transform coding effective for images and videos*

- *Intuition on why DCT/DWT leads to sparse representations (spatial and temporal correlations, etc?)*

- *Explain MSE, PSNR*

An important choice in any data analysis is how to represent the data. In this chapter we will introduce two particular ways of representing data, the *Discrete Cosine Transform* (DCT) and the *Discrete Wavelet Transform* (DWT). Applying such a transform to a raw a signal in the time or space domain results in a representation of that same signal in an alternative domain. The focus is on DWT and DCT since, when transforming image and video data, these two transforms tend to output sparse (or near-sparse) representations of the data.

The main objective of this chapter is to explain how to construct the basis matrix for the DCT and the DWT.

3.1 Discrete Cosine Transform

An aside on the DCT. Used in old JPEG standard (ref). Formulae. Interpretations. Pictures.

Fig. 3.1 Original image $\boldsymbol{x}$ (left) and its Haar basis transformation $\boldsymbol{w}$ (right). See next chapter for more details on Haar wavelets.

3.2 Discrete Wavelet Transform

3.2.1 Haar wavelets

Finding a set of basis functions $\boldsymbol{\Psi}$ that achieve such a transformation lies at the heart of many lossy compression techniques. For instance, in image processing the JPEG 2000 standard is a widely used lossy compression technique that relies on this principle. The original image $\boldsymbol{x}$ is transformed into $\boldsymbol{w}$ using the so-called *Discrete Cosine Transform*. The basis matrix $\boldsymbol{\Psi}$ is orthogonal, so $\boldsymbol{x}$ and $\boldsymbol{w}$ have the same l_2 norm. In the original signal $\boldsymbol{x}$ the length is spread across many of its coefficients. On the other hand, most of the length of $\boldsymbol{w}$ is concentrated in a few of its coefficients. A large fraction of the entries in $\boldsymbol{w}$ are very close to zero. By deleting these entries in $\boldsymbol{w}$ and only storing the non-zero coefficients (and the corresponding basis functions), we can obtain a compressed version $\hat{\boldsymbol{w}}$. This allows us to significantly reduce the amount of data that needs to be stored without affecting the visual quality in the reconstructed image $\hat{\boldsymbol{x}} = \boldsymbol{\Psi}\hat{\boldsymbol{w}}$.

It is important to note here that the choice of basis functions $\boldsymbol{\Psi}$ typically has a significant effect on the performance of the reconstruction algorithms.

The simplest wavelet basis transformation is based on the Haar wavelets. Figure 3.1 vizualises a basis transformation of the original image $\boldsymbol{x}$ to $\boldsymbol{w}$. This example uses a Haar wavelet basis at the first scale. We will explain the Haar wavelet transformations of images and videos in more detail in the next chapter. Dark areas correspond to small coefficients. Note that most entries in $\boldsymbol{w}$ are near zero. In practice, we approximate these entries as zero and treat $\boldsymbol{w}$ as sparse.

3.2.2 Daubechies Wavelets

Intuition. Where do the coeffs come from. Matrices. Boundary conditions.

3.3 Forming the basis matrix

1D, 2D, 3D case. Different scales

For a deeper introduction into wavelets see [6]. For more information on wavelet compression techniques, see [3].

References

[1] Baraniuk, R. G. (2007). Compressive sensing. *IEEE signal processing magazine*, 24(4).

[2] Candès, E. J. and Wakin, M. B. (2008). An introduction to compressive sampling. *Signal Processing Magazine, IEEE*, 25(2):21–30.

[3] DeVore, R. A., Jawerth, B., and Lucier, B. J. (1992). Image compression through wavelet transform coding. *Information Theory, IEEE Transactions on*, 38(2):719–746.

[4] Pilikos, G. (2014). Signal reconstruction using compressive sensing. MPhil thesis, University of Cambridge.

[5] Shannon, C. E. (1949). Communication in the presence of noise. *Proceedings of the IRE*, 37(1):10–21.

[6] Stollnitz, E. J., DeRose, T. D., and Salesin, D. H. (1995). Wavelets for computer graphics: a primer. 1. *Computer Graphics and Applications, IEEE*, 15(3):76–84.

[7] Taubman, D. and Marcellin, M. (2012). *JPEG2000 Image Compression Fundamentals, Standards and Practice: Image Compression Fundamentals, Standards and Practice.* The Springer International Series in Engineering and Computer Science. Springer US.

[8] Tipping, A. and Faul, A. (2002). Analysis of sparse bayesian learning. *Advances in neural information processing systems*, 14:383–389.

[9] Tipping, M. E. (2001). Sparse bayesian learning and the relevance vector machine. *The journal of machine learning research*, 1:211–244.

[10] Tipping, M. E., Faul, A. C., et al. (2003). Fast marginal likelihood maximisation for sparse bayesian models. In *AISTATS*.