A Short Overview of Prediction Methods for Image Compression

Luka Lukač[1], Aljaž Jeromel[1], Libor Váša[2]

[1]*Faculty of Electrical Engineering and Computer Science, University of Maribor*
[2]*Department of Computer Science and Engineering, University of West Bohemia*
E-mail: luka.lukac@um.si

Abstract

In this paper, an overview of prediction methods used in image compression is presented. The overview contains the state-of-the-art methods, including PNG, JPEG, JPEG-LS, and WebP, and some newer methods that could potentially become state-of-the-art in the future. While the state-of-the-art methods mostly rely on fixed prediction schemes, newer methods explore the use of genetic algorithms, artificial neural networks and optimisation approaches. The aim of this paper is to provide a short description of recent advances in this field.

1 Introduction

Data compression is one of the most important topics in computer science [1]. Even though the storage capacities of new computers are much larger than in the past, data transfer speeds are still relatively slow due to bandwidth limitations. Therefore, data compression is still a very important topic of research, as it allows faster data transfers. This is especially evident on social networks, where huge amounts of data in forms of images or video are being transferred each second.

Images are a special kind of data. They are usually represented as a two-dimensional matrix of pixels, where each pixel is represented by either one bit (for bi-level images), one byte (for grayscale images), three bytes (for colour images without transparency), or four bytes (for colour images with transparency). Because the images are two-dimensional, it makes sense that the neighbouring pixels are in some way correlated. This is the main idea behind prediction-based image compression. The value of current pixel to be encoded is inferred (predicted) from a set of already-encoded (and decoded on the decoder side) neighbouring pixels, and only the difference (the residual) between the predicted value and the actual value of the pixel are encoded. This approach reduces the entropy of data, as the spatial correlations between pixels are reduced. A good prediction therefore yields efficient compression. In an extreme case, when the predictions exactly match the encoded values, the sequence of residuals reduces to a sequence of zeroes with zero entropy. In practice, this state is of course never achieved, however, the accuracy of the prediction is directly linked with the compression efficiency.

Throughout the years, a lot of prediction methods have been developed. The aim of this paper is to provide a short overview of the state-of-the-art methods, as well as recently developed methods. This overview could help the readers pick the best prediction method for their algorithm, or suggest some ideas where the prediction could be improved. The rest of this paper is structured as follows. An overview of prediction methods is given in Section 2. Section 3 contains the conclusion.

2 Prediction Methods

Lossless, as well as some lossy compression methods are based on reducing the statistic redundancy in data [2]. The majority of techniques dealing with this issue utilise the concept of the current pixel prediction according to its neighbourhood. Most of the time, neighbouring pixels have similar values, therefore, only the difference between the prediction value and the actual current pixel value is encoded. Consequently, better compression results are achieved, as spatial redundancy is removed [3]. An example of a prediction neighbourhood is depicted in Figure 1, where the current pixel is painted grey and marked with X, while the neighbouring pixels are marked with letters A–D. Note that the values at these pixels are already known to the decoder if the image is processed row by row, left to right, and therefore, their true (corrected) values can be used to derive the prediction.

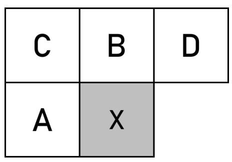

Figure 1: Prediction neighbourhood, where the current pixel is marked with an X.

This Section contains an overview of prediction methods. A survey containing mathematical discussion and definitions on prediction is given in [4].

2.1 PNG

Portable Network Graphics (PNG) [5] uses five different predictors, also known as filters, which are applied separately by the encoder for each row of pixels in the image. $None$ does not give a prediction, Sub predicts the value of the pixel X by the value of the pixel A (Figure 1), Up predicts the value of the pixel above the current one, while $Average$ takes the mean value between pixel's neighbours at the top and to the left. Finally, $Paeth$ predicts the value according to values $P = |B - C|$, $Q = |A - C|$ and $R = |A + B - 2C|$. If P is the smallest among of the three, left pixel value is predicted. If Q is the minimal value, the upper pixel value is predicted. In cases when R is the smallest value, the value of the left upper pixel is given as a prediction. The filters, along with their predictions, are summarised in Table 1.

Table 1: PNG format predictor types.

Name	Prediction
$None$	0
Sub	A
Up	B
$Average$	$\lfloor \frac{A+B}{2} \rfloor$
$Paeth$	A or B or C

2.2 JBIG

Joint Bi-Level Image Experts Group (JBIG) is a format for compressing binary image data [6]. In the standard, three different algorithms for pixel prediction are defined:

- Differential Layer Typical Prediction (DTP), where 3x3 neighbourhood of the current pixel is observed, assuming presence of a lower resolution image. If all pixels are the same colour, that colour is predicted, which is true for most cases. The correctness of the prediction is encoded in a pseudo-pixel (a structure indicating the metadata of the prediction).
- Base Layer Typical Prediction (BTP), which is considered when the lowest resolution layer is encoded. Two consecutive lines are compared, and if they are the same, a pseudo-pixel is encoded carrying that information. Consequently, the encoding of the second line can be skipped.
- Deterministic Prediction (DP), where redundant information between a low resolution and a high resolution image is removed. For example, if a black high resolution pixel that belongs to a white low resolution pixel is decoded, other pixels belonging to the area of a white low resolution pixel are predicted to be white.

2.3 Lossless JPEG

In contrast with PNG, lossless format of Joint Photographic Experts Group (lossless JPEG) uses 8 different predictors [7]. The first three predictors are the same as PNG's, as no prediction is made at selection 0, left pixel is predicted at selection 1 and upper pixel is predicted if stumbled upon selection 2. Selection 3 predicts the value of the left upper pixel, while linear combinations of neighbouring pixels A, B and C are predicted at selections from 4 to 7. During the encoding process, any of the predictors, depicted in Table 2, can be used (selected separately for each pixel).

Table 2: Lossless JPEG predictor types.

Selection	Prediction
0	/
1	A
2	B
3	C
4	$A + B - C$
5	$A + ((B - C)/2)$
6	$B + ((A - C)/2)$
7	$(A + B)/2$

2.4 JPEG-LS

Unlike some other methods, JPEG-LS encoding method selects the type of the prediction according to several conditions in the prediction neighbourhood [8]. To improve prediction quality, Median Edge Detection predictor is incorporated into the standard. Its basic principle enables edge detection, which can, consequently, yield better prediction accuracy. When a diagonal edge is indicated by the edge predictor, the prediction is given by to value of C: if brighter than A and B, the current pixel probably should be darker than A and B, and vice versa. In other cases, a linear combination of A, B, and C is given as a prediction. The JPEG-LS predictor is summarised in Equation 1.

$$X = \begin{cases} min(A, B); & C > max(A, B) \\ max(A, B); & C < min(A, B) \\ A + B - C; & \text{otherwise} \end{cases} \tag{1}$$

2.5 JPEG XR

JPEG XR introduces an adaptive prediction, which occurs solely if two neighbouring macroblocks (basic data unit with Direct Current (DC), Low Pass (LP) and High Pass (HP) components, consisting of 16x16 pixels) are similar enough [9]. The direction of the macroblock, upon which the prediction is made, depends on the similarity of macroblocks to the left, top and top-left, thus the predictor is called adaptive. DC and LP coefficients are predicted between two macroblocks while HP coefficients are predicted within each macroblock.

2.6 JPEG XL

Similarly to JPEG XR, JPEG XL also utilises the concept of adaptive prediction [10]. In lossy mode, the predictor chooses between eight prediction modes according to the prediction neighbourhood. On the other hand, in lossless mode, the number of prediction modes is reduced to four. Furthermore, instead of choosing one of the predictors, the encoder computes a weighted average of all four possibilities, which is considered a prediction value.

2.7 WebP

WebP was introduced by Google with the intention of optimising image loading times on the Web while retaining their quality [11]. Although based on JPEG, WebP uses different prediction methods, borrowed from video compression standard VP8. Both standards implement four prediction modes, used with 4x4, 8x8 and 16x16 macroblocks. The leftmost column of the block is denoted by L, and the topmost row of the block is denoted by A. The standard allows the following prediction modes:

- H_PRED (horizontal prediction), where each column of the prediction block is a copy of L,
- V_PRED (vertical prediction), where each row of the prediction block is a copy of A,
- DC_PRED (DC prediction), where the block is filled with a single value, calculated as the average value of the pixels in the row above A and pixels left to L, and
- TM_PRED (TrueMotion prediction), where the block is filled with values of A and L for each pixel with coordinates x, y. After that, predictions for each pixel are calculated as the difference between the average value of the column $x-1$ and the row $y-1$, and the value of the pixel at $x-1, y-1$.

2.8 CALIC

CALIC is a context-based image compression algorithm, presented in 1996 by Wu and Memon [12]. It uses a Gradient Adjusted Predictor, GAP, which detects the direction of the edge at the current pixel and uses thresholding to classify it as none, weak, normal, or sharp edge. The pixel value is then predicted according to its $5x5$ neighbourhood and its classification. Based on the prediction error on the next pixel, the prediction is updated by error quantisation and context modelling, which results in a context-based adaptive prediction.

2.9 TMW

In 1997, Meyer and Tischer presented a lossless image compression algorithm, TMW [13], based on a combination of different linear predictors. TMW combines three types of predictors:

- pixel predictors, where the predicted value of the current pixel is calculated by a weighted sum of reference pixels,
- sigma predictors, which predict the pixel predictor's error, and
- blending predictors, which predict the suitability of pixel predictors for the current pixel based on its neighbourhood.

TMW first performs an image analysis, where the weights for linear pixel predictors are calculated. This is done by prediction of errors and suitability of each of the pixel predictors. TMW's uniqueness comes from the fact that each pixel can be predicted using more than one linear predictor.

2.10 Edge-Directed Prediction

In 2001, Li and Orchard presented an edge-directed prediction method for lossless image compression [14]. The method uses a so-called training window to predict the current pixel. For an arbitrary natural number T, the window contains $2T(T+1)$ samples. The window is composed of two connected rectangular parts, as seen in Figure 2. Then, the Least-Squares (LS) optimisation [15] is used to calculate the optimal prediction coefficients for the image.

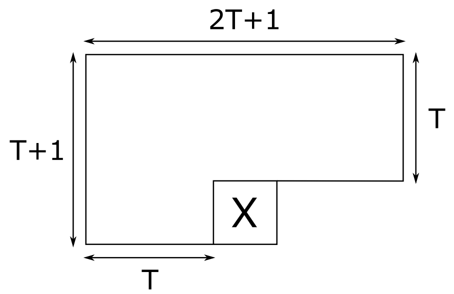

Figure 2: Moving window used for prediction.

2.11 Prediction Using Genetic Programming

In 2009, Takamura, Matsumura, and Yashima presented a genetic algorithm-driven method for prediction in lossless image compression [16]. The image is analysed first, and a general predictor is expressed as a tree structure, where the leaf nodes represent the features (neighbouring pixel values, detected edges, results of different prediction methods, etc.), and non-leaf nodes represent conditional branching or mathematical operations on those features. This predictor design is then put through a genetic algorithm, which attempts to optimise it according to the image being compressed. The pixel neighbourhood used for prediction is the same as the one in Figure 2, where $T \in \{1, 2\}$.

2.12 Prediction Using Neural Networks

With the rapid development of convolutional neural networks (CNN) in the recent years, prediction methods are no longer limited to purely algorithmic schemes. In 2021, Yang et al. have presented a transformer-based neural network for pixel prediction [17]. The network receives an image, which is sent through a pre-trained ResNet architecture. The result is linearly projected (transformed from 2D to 1D) and sent into a sequence of transformer layers. The output of those layers is then passed to an attention gate, and upsampled to obtain the final prediction. In the paper, prediction was used to estimate depth of the image at each pixel.

Also in 2021, Rhee et al. presented a method for lossless image compression using CNN, where the duplex neural networks are used to predict both the pixel value and the prediction error [18]. The method firstly transforms the RGB colour space to YUV. Then, the duplex neural network predicts sequentially the value and context for each pixel in raster scan order from the neighbouring pixels. First, the Y colour channel is predicted, and the result is used in further prediction of U and V colour channels and context of the pixel. The probability distribution of prediction errors is then used for compression with the adaptive arithmetic coder.

In 2022, Guo, Zhang, Feng, and Chen proposed a method for lossy image compression, based on CNN [19]. The two known pixels, closest to the current pixel (pixels A and B in Figure 1), are used for context modelling. Then, the value of the current pixel is predicted by feeding into a CNN the 5x5 neighbourhood around the current pixel, the values of the previously decoded colour channels, and the aforementioned context.

3 Conclusion

Prediction is one of the crucial components of many image compression techniques, which can severely enhance compression ratio while retaining the original quality of an image. Exploiting similarity between pixels in a close neighbourhood, redundant information does not need to be encoded, leading to large reduction in data rate. Up to this date, many different prediction methods have been developed. The simplest routines usually give predictions according to values of neighbouring pixels and their linear combinations. Although simple in nature, a lot of state-of-the-art image compression formats, such as PNG, JBIG, and lossless JPEG, successfully utilise such approaches. Another, more sophisticated approach is based on edge detection, enabling more accurate predictions in areas that feature prominent oscillations of pixel values. There are plenty of other methods for prediction, such as adaptive prediction between the macroblocks, motion prediction, genetic algorithms, and neural networks. Although not so widely used, in some cases, they can produce even better results than conventional prediction techniques.

Acknowledgement

This research was funded by the Slovene Research Agency under Research Project J2- 4458 and Research Programme P2-0041, and the Czech Science Foundation under Research Project 23-04622L.

References

[1] D. Salomon, G. Motta: Data Compression: The Complete Reference, 5th Edition. Springer, 2010.

[2] M.A. Rahman, M. Hamada, J. Shin: The Impact of State-of-the-Art Techniques for Lossless Still Image Compression. Electronics, vol. 10, no. 3, pp. 60, 2021.

[3] A. Avramović, B. Reljin: Gradient Edge Detection Predictor for Image Lossless Compression. Proceedings ELMAR-2010, Zadar, Croatia, pp. 131-134, 2010.

[4] N. Merhav: Universal Prediction. IEEE Transactions on Information Theory, vol. 44, no. 6., pp. 2124-2147, 1998.

[5] G. Roelofs: PNG: The Definitive Guide. O'Reilly & Associates, Inc., 1999.

[6] C.M. Smith: Efficient Software Implementation of the JBIG Compression Standard. Thesis, Rochester Institute of Technology, 1993.

[7] W.B. Pennebaker, J.L. Mitchell: JPEG: Still Image Data Compression Standard. Springer Science & Business Media, 1992.

[8] N. D. Memon, X. Wu, V. Sippy, G. Miller: Interband Coding Extension of the New Lossless JPEG Standard. Visual Communications and Image Processing, vol. 3024, pp. 47-58, January 1997.

[9] S.S. Jadhav, S.K. Jadhav: JPEG XR an Image Coding Standard. International Journal of Computer and Electrical Engineering, vol. 4, no. 2, pp. 137, 2012.

[10] J. Alakuijala, R. van Asseldonk, S. Boukortt, M. Bruse, I.-M. Comșa, M. Firsching, T. Fischbacher et al.: JPEG XL Next-Generation Image Compression Architecture and Coding Tools. Applications of Digital Image Processing XLII, vol. 11137, pp. 112-124., 2019.

[11] Google Developers: Compression Techniques — WebP — Google for Developers. https://developers.google.com/speed/webp/docs/compression (Access: 7. 7. 2023).

[12] X. Wu, N. Memon: CALIC - A Context Based Adaptive Lossless Image Codec. In Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 4, pp. 1890-1893, 1996.

[13] B. Meyer, P. Tischer: TMW - a New Method for Lossless Image Compression. ITG FACHBERICHT, vol. 2, pp. 533-540, 1997.

[14] X. Li, M. T. Orchard: Edge-Directed Prediction for Lossless Compression of Natural Images. IEEE Transactions on Image Processing, vol. 10, no. 6, pp. 813-817, June 2001.

[15] D. D. Morrison: Optimization by LeastSquares. SIAM Journal on Numerical Analysis, vol. 5, no. 1, pp. 83-88, 1968.

[16] S. Takamura, M. Matsumura, Y. Yashima: A Study on an Evolutionary Pixel Predictor and Its Properties. In Proceedings of 16th IEEE International Conference on Image Processing (ICIP), pp. 1921-1924, 2009.

[17] G. Yang, H. Tang, M. Ding, N. Sebe, E. Ricci: Transformer-Based Attention Networks for Continuous Pixel-Wise Prediction. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 16269–16279, 2021.

[18] H. Rhee, Y. I. Jang, S. Kim, N. I. Cho: Lossless Image Compression by Joint Prediction of Pixel and Context Using Duplex Neural Networks. IEEE Access, vol. 9, pp. 86632-86645, 2021.

[19] Z. Guo, Z. Zhang, R. Feng, Z. Chen: Causal Contextual Prediction for Learned Image Compression. IEEE Transactions on Circuits and Systems for Video Technology, vol. 32, no. 4, pp. 2329-2341, April 2022.