Clustering-based quantisation for PDE-based image compression

Laurent Hoeltgen · Pascal Peter · Michael Breuß

Abstract Optimal known pixel data for inpainting in compression codecs based on partial differential equations is real-valued and thereby expensive to store. Thus, quantisation is required for efficient encoding. In this paper, we interpret the quantisation step as a clustering problem. Due to the global impact of each known pixel and correlations between spatial and tonal data, we investigate the central question, which kind of feature vectors should be used for clustering with popular strategies such as $k$-means. Our findings show that the number of colours can be reduced significantly without impacting the reconstruction quality. Surprisingly, these benefits are negated by an increased coding cost in compression applications.

Keywords Laplace interpolation · Inpainting · Compression · Quantisation · Clustering · Partial differential equations

Mathematics Subject Classification 65N99 · 90C25 · 97N50

1 Introduction

A major challenge in data analysis is the reconstruction of a signal from very few data points. In image processing this interpolation problem is called inpainting. Recent advances have shown that accurate reconstructions from a small sample of well-chosen image points are possible by using methods based on partial differential equations (PDEs) [1,4,6,11,18,21,24,32]. These efforts have also been exploited for lossy image compression schemes that can nowadays compete with other state-of-the-art approaches [29,32]. Thus, these methods provide a viable alternative to classic transform-based approaches, such as JPEG and JPEG 2000, as well as adaptive schemes such as in [26]. Even simple methods that rely on the interpolation capabilities of the Laplacian may yield more accurate reconstructions than JPEG 2000 [27]. Unfortunately, the underlying optimisation problem concerned with selecting optimal point locations is computationally intensive. In addition, the resulting data are often difficult to store efficiently. The pixel locations may be scattered arbitrarily inside the image domain, and the corresponding colour values are often computed in floating point precision. Storing these data as-is is prohibitively expensive and often unnecessary. The human visual system is only capable of distinguishing about thirty different shades of grey [20]. These thirty colours can be encoded with 5 bit instead of the 64 bit required for a floating point value in double precision. Finding the optimal number and distribution of these grey values is however a complicated task. Nevertheless, it is reasonable to expect that small changes to the colour values will only yield small changes in the reconstruction. This assumption allows us to replace real-valued colours by their closest integer-valued neighbours without causing to much damage.
This work extends the preliminary findings from [16]. There, the authors investigated clustering schemes and quality assessment measures to provide fast strategies to determine a good number of quantised colours. In [16], the clustering was not evaluated against the compression ratio of the final compressed file but only against the reconstruction quality of the clustered data. Nevertheless, the results indicated that simple clustering approaches yield good quantisations of the co-domain. In addition, the authors showed that these clusterings can be evaluated with quality measures such as the silhouette coefficient.

Our Contribution In this paper we investigate how well different clustering optimisation techniques fare in a data compression context. Thus, we extend the results from [16] by providing a similar evaluation on the compression ratio. In addition, we rank all considered clustering strategies with respect to their effectiveness and show that the quantisation must be done with the compression ratio in mind and not with respect to the reconstruction quality.

To this end we will discuss several algorithms as well as quality measures and show that not only the number of clusters, but also the final distribution of the labels is important. The distribution of the labels is mostly influenced by the underlying clustering algorithm, whereas the number of clusters can be optimised independently for any strategy. The latter task is also considered in detail. Several strategies are proposed. This discussion extends the findings from [16] where only the silhouette coefficient was taken into consideration.

2 Partial differential equation-based image compression

Even though most diffusion-type partial differential equations can be used to reconstruct the data, previous works have shown that homogeneous, linear diffusion ranks among the best performing ones [27] in the compression context. We also refer to [2] for a more general overview on PDE-based inpainting methods. The simplicity of homogeneous diffusion inpainting allows a thorough mathematical analysis and an extensive optimisation [18,19]. Since our sought optimal quantisation is likely to depend strongly on the underlying reconstruction process, we focus exclusively on homogeneous diffusion in this work. Nevertheless, many of the presented strategies can be extended to other inpainting methods.

A good image compression approach consists of several components. Besides the reconstruction method it is also necessary to optimise the considered data and to design an efficient encoding. Let us now briefly review the involved reconstruction and optimisation steps in the next paragraphs.

2.1 Inpainting with homogeneous diffusion

Inpainting with homogeneous diffusion (sometimes also called Laplace interpolation) is a rather simple reconstruction method that is well suited for highly scattered data in arbitrary dimensional settings. It can be modelled as follows. Let \( f : \Omega \rightarrow \mathbb{R} \) be a smooth function on some bounded domain \( \Omega \subseteq \mathbb{R}^2 \) with a sufficiently regular boundary \( \partial \Omega \). Throughout this work, we will restrict ourselves to the case of grey value images, even though many of the results hold for arbitrary numbers of colour channels. Moreover, let us assume that there exists a closed nonempty set of known data \( \Omega_K \subseteq \Omega \) that will be interpolated by the underlying diffusion process. Homogeneous diffusion inpainting considers the following partial differential equation with mixed boundary conditions:

\[
-\Delta u = 0 \quad \text{on } \Omega \setminus \Omega_K
\]

\[
\text{with } \begin{cases} 
  u = f & \text{on } \partial \Omega_K \\
  \partial_n u = 0 & \text{on } \partial \Omega \setminus \partial \Omega_K
\end{cases}
\]

where \( \Delta \) represents the Laplacian differential operator; \( \partial_n u \) denotes the derivative of \( u \) in the outer normal direction along the image boundary \( \partial \Omega \setminus \partial \Omega_K \). Also, \( \Omega \setminus \Omega_K \) describes the region to be reconstructed (see Fig. 1 for a sketch of the set-up). We assume that both boundary sets \( \partial \Omega_K \) and \( \partial \Omega \setminus \partial \Omega_K \) are nonempty. Equations of this type are commonly referred to as mixed boundary value problems and sometimes also as Zaremba’s problem [37]. The existence and uniqueness of solutions has been studied during the last century. Showing that (1) is indeed solvable is by no means a trivial feat. We refer to [13] for an extensive study of linear elliptic partial differential equations and to [17] for a more particular analysis of (1). A particularly simple case is the 1-D setting, where the solution can obviously be expressed using piecewise linear splines interpolating data on \( \partial \Omega_K \).

![Fig. 1 Example set-up as considered in (1). The grey regions are known data, and the white parts need to be recovered](image-url)
2.2 Optimal inpainting positions

Finding good inpainting positions is a task related to the free knot problem from interpolation theory [3,15]. In the special context of image inpainting and image compression the authors of [6,8,18,24] suggested several strategies to find good locations. Mainberger et al. [24] propose stochastic optimisations that bear similarities to simulated annealing approaches. Chen et al. [6] use fast bi-level optimisation schemes, and in [18] Hoeltgen et al. present an optimal control model. Finally, the works of Demaret et al. [8,9] use adaptive triangulations and mesh optimisation strategies, whereas Ochs et al. [25] suggest fast numerics for the task at hand.

2.3 Continuous grey value optimisation

In [19,24] the authors discuss approaches to find the best pixel values. These algorithms have in common that they all yield real-valued floating point results. If we denote the reconstruction operator that solves (1) for a given mask \( \Omega_K \) by \( M(\Omega_K) \), then this problem can be formulated as

\[
\arg\min_u \left\{ \| M(\Omega_K)u - f \|^2 \right\}
\] (2)

Since our initial PDE is linear, the task stated in (2) corresponds to a linear least squares problem and can be solved efficiently. An alternative discrete optimisation has been suggested by Schmaltz et al. [32]. The latter approach iterates over the inpainting data and selects in a greedy way the currently best quantised grey value. The iterations continue until a convergent state is reached. While this approach comes conceptually close to clustering strategies, it has two notable drawbacks: (i) The approach requires that the distribution of the quantised values is fixed at the beginning, i.e. it does not change during the iterations; (ii) to test a certain grey value for its quality requires the solution of a large and sparse linear system, i.e. the computational costs are considerable as the algorithm requires the solution of thousands of linear systems. Nevertheless, we remark that it yields excellent qualitative results.

3 Clustering inpainting data

The following paragraphs discuss potential choices for the feature selection as well as for the algorithmic execution. Our motivation is to use models that come conceptually close to the continuous grey value optimisation and the quantisation-aware approach of Schmaltz et al. [32]. Therefore, we focus on the squared Euclidean distance as cost function and seek the best set of grey values for the mask pixel locations. However, we will refrain from using the reconstruction inside our algorithms in order to obtain fast methods.

3.1 Feature selection

Our set-up provides us the full image data \( f \) as well as an optimised inpainting mask \( \Omega_K \). There exist many possible choices to extract interesting features from these data. The following eight feature choices seem reasonable:

1. Position and value of all image pixels
2. Position and value of all mask pixels
3. Grey values of all image pixels (i.e. discarding the positional information)
4. Grey values of all mask pixels
5. The histogram of all image pixels
6. The histogram of all mask pixels
7. The colour map (i.e. histogram without frequency count) of all image pixels
8. The colour map of all mask pixels

The following observations can be made with respect to the feature choices.

- We remark that including the pixel position into the feature vector renders each sample unique. In that case it makes no sense to track their frequency as it will always be 1.
- Approaches that exploit the full image information have better chances at adapting to features in the image that are not covered by the mask pixels alone. On the other hand, optimised positions for the inpainting with homogeneous diffusion are usually in the vicinity of edges and other important structures. Therefore, it is likely that the image information provided by the mask pixels is already sufficient to determine a good clustering. In addition, for compression purposes the inpainting masks are usually very sparse. Thus, we also benefit from a smaller feature set and a higher performance.
- Strategies that focus exclusively on the grey values of an image are likely to perform fastest since they can be carried out on the histogram. As soon as we include the spatial position into the feature set we obtain as many unique feature values as pixels. For large images we may encounter memory or run time restrictions.

Finally, let us remark that from a clustering perspective it is irrelevant whether we consider grey values or colour images. Only the mask optimisation and inpainting steps are more difficult to carry out for colour images.
3.2 Clustering approaches for image quantisation

In this work we opt for some of the most popular and best understood approaches to classify our data. Since clustering approaches are commonly subdivided into partitional and hierarchical methods, we select the following representatives:

1. **partitional clustering** We use the $k$-means++ variant of the venerable $k$-means algorithm of Lloyd [22].

2. **hierarchical clustering** We employ a bottom-up strategy where initially each feature represents a single cluster. These clusters are merged step-by-step using Ward’s criterion [36].

3. **probabilistic clustering** As an alternative to the “hard” labelling performed by the $k$-means approach we also consider a “softer” variant using a Gaussian mixture model and employ the popular expectation maximisation algorithm of Sundberg [33,34] to compute the probabilities that a feature belongs to a certain cluster. Probabilistic clustering can also be seen as a form of fuzzy clustering, see e.g. [12].

We also refer to [14] for a discussion on the viability of various clustering techniques.

3.3 Quality measures and optimal number of clusters

When evaluating a partitioning of the data we need to consider several criteria. First, the quality of the clustering itself: Features in the same group should indeed be similar and yet also distinct from observations from other clusters. For our purposes it is also important to assess the quality of the reconstruction as well as the compression ratio of the final file. In particular, the latter quantity is difficult to estimate. Our quantised data should have few different grey values (i.e. a small number of clusters) such that the file compression can be efficient. However, it is also essential that the reconstruction quality is fair, which is likely to require many different grey values (i.e. a large number of clusters) in floating point precision. Thus, there is a trade-off between reconstruction quality and compression ratio. This influence can be steered by optimising the number of clusters. From the large pool of quality measures available in the literature ([10] lists more than 50) we chose the following to help us identify an optimal grouping of our data. In this description, $k$ denotes always the number of clusters and $N$ the number of features.

1. The Calinski–Harabasz (CH) criterion [5], also called variance ratio criterion, considers the quantity

$$
\frac{\sum_{j=1}^{k} \|m_j - m\|^2}{\sum_{j=1}^{k} \sum_{x \in C_j} \|x - m\|^2} \frac{N - k}{k - 1}
$$

(3)

where $m_j$ is the centroid of the cluster $C_j$ and $m$ the overall mean of all features. Higher values indicate better clusterings.

2. The Davies–Bouldin (DB) criterion [7] is based on a ratio of within-cluster and between-cluster distances. The Davies–Bouldin index is defined as

$$
\frac{1}{k} \sum_{j=1}^{k} \max_{i \neq j} \left\{ \frac{\bar{d}_j + \bar{d}_i}{\max\{d_{i,j}, d_{i,s}\}} \right\}
$$

(4)

with $\bar{d}_i$ being the average distance of the centroid to each element in the cluster $C_i$ and where $d_{i,s}$ is the distance between the centroids of the clusters $C_s$ and $C_i$. Lower values are better.

3. The GAP criterion [35] formalises the well-known L-term heuristic by estimating the “elbow” location. The “elbow” occurs at the most dramatic decrease in error measurement, i.e. the largest gap value. To this end, the within-cluster sum of squares is compared to its expectation under an appropriate null reference distribution of the data. This leads to the following formula for the GAP value:

$$
\text{GAP}_N(k) := E_N^* [\log(W_k)] - \log(W_k)
$$

(5)

where $E_N^*$ denotes the expectation under a sample of size $N$ from the reference distribution. The optimal number of clusters is given by $k$ which maximises the GAP value.

4. The Silhouette index [30] of a cluster $C$ is defined to be the average silhouette value of all its elements. The silhouette value of a single element $x \in C$ is defined as

$$
\begin{cases}
0, & \text{if } \text{dist}(x, C) = 0 \\
\frac{\text{dist}(x, B) - \text{dist}(x, C)}{\max\{\text{dist}(x, C), \text{dist}(x, B)\}}, & \text{else}
\end{cases}
$$

(6)

where $\text{dist}(x, C)$ is the average distance of $x$ to all other elements from the cluster $C$ and where $B$ is the next closest cluster (found by minimising the average distance between $x$ and all elements from the cluster $B$). Comparing the silhouette indices for all possible clusters gives an estimate on how well structured the clustering is. These indices should be as close to 1 as possible.

We remark that the GAP criterion is one of the most flexible approaches but also the most expensive measure in terms of computational effort.

 Springer
4 Numerical experiments

4.1 Performance of the clustering algorithms

In total the proposed 8 features from Sect. 3.1, 3 algorithms from Sect. 3.2 and 4 quality criteria from Sect. 3.3 allow 96 possible combinations to be evaluated. An exhaustive testing would require the evaluation of all these approaches on several test images, thus increasing the data to be analysed even more. Certainly not all of these combinations are reasonable, and some strategies also require considerable run times such that they become impractical for most applications. Thus, we have decided to restrict ourselves to the following choices. This list is also motivated by the findings from [16] where it has been noticed that the $k$-means and hierarchical clustering are more reliable and usually perform better than the probabilistic clustering with a Gaussian mixture model.

1. $k$-means on the position and value of all mask pixels
2. hierar. clustering on position and value of all mask pixels
3. probab. clustering on position and value of all mask pixels
4. $k$-means on the pixel values of all mask pixels
5. hierar. clustering on the pixel values of all mask pixels
6. $k$-means on the histogram of all mask pixels
7. $k$-means on the histogram of all image pixels
8. $k$-means on the unique grey values of all mask pixels
9. $k$-means on the unique grey values of all image pixels

Our experiments were carried out on the data set presented in Fig. 2. The considered test image is of size 256 $\times$ 256 and has 170 unique grey values. These are bytewise coded (i.e. the maximal amount of different grey values is 256), and the corresponding binary mask has a density of 5%. It has been obtained with the optimal control approach of Hoeltgen et al. [18]. The histograms of the image and of the pixels indicated by the mask are depicted in Fig. 3. The considered test image and its histogram are representative for a large class of images. The image has smooth regions as well as textured areas. The corresponding histograms are also very generic. They do not show any particular patterns. As such, we believe that the presented findings would be similar for other images. The methods presented in this work are also agnostic towards how the mask was obtained.

Let us in a first step evaluate our considered criteria on the trui test image from Fig. 2. The results of our experiments are listed in Table 1. We see that integrating the position into the feature vectors significantly deteriorates the reconstruction quality (rows 1–3 in Table 1). The suggested optimal numbers of clusters however look reasonable, except for the GAP criterion which tends to suggest rather large numbers of clusters. The other experiments (rows 4–9 in Table 1) yield, at least in terms of mean squared error, good results. Also here the optimal number of clusters is often overestimated. Notable exceptions are the GAP criterion for the set-ups 4, 5, 8 and 9 where the suggested number of clusters was around 12 and the reconstruction error around 52. For the set-up 6, the GAP criterion suggested 36 colours with a corresponding mean squared error of 48.91. Considering that the reconstruction with the original 170 colours has an mean squared error of 46.96, this is actually a very satisfying result. Also the suggested number of clusters coincides with our expectations.

Figure 4 plots the evolution of the reconstruction error as a function of the number of clusters for the $k$-means method as well as for the hierarchical clustering. The $k$-means method performs slightly better than the hierarchical scheme. For 35 or more clusters, there is hardly any improvement in the reconstruction anymore and the errors converge towards the mean squared error of the original data. The steady and rather stable decrease of the error might explain why most clustering methods tend to overestimate the number of clusters. In terms of error, the findings for 40 clusters are almost as good as those with 72 clusters. Without an explicit requirement that the optimal number of clusters should be small, these results are difficult to discern.
Table 1  Clustering results for the true test image. We tested all cluster sizes ranging from 12 to 72

| Model | Mean squared error | Best number of clusters |
|-------|--------------------|-------------------------|
|       | Silhouette | DB | CH | GAP | Silhouette | DB | CH | GAP |
| 1     |  404     |  246 |  350 |  125 |  16     |  25 |  15 |  66 |
| 2     |  380     |  199 |  271 |  147 |  13     |  38 |  27 |  68 |
| 3     |  418     |  316 |  295 |  152 |  12     |  15 |  19 |  71 |
| 4     | 47.15    | 47.08 | 46.95 | 54.88 |  63     |  72 |  70 |  12 |
| 5     | 47.26    | 47.32 | 47.26 | 52.63 |  72     |  69 |  72 |  14 |
| 6     | 63.17    | 47.51 | 47.66 | 48.91 |  12     |  72 |  72 |  36 |
| 7     | 52.50    | 51.65 | 53.01 | 54.86 |  72     |  72 |  72 |  66 |
| 8     | 51.30    | 46.63 | 46.86 | 50.34 |  12     |  63 |  69 |  12 |
| 9     | 52.94    | 46.57 | 46.62 | 49.88 |  12     |  72 |  67 |  12 |

The mean squared error is always given for the corresponding suggested optimal number of clusters. As we can see, most methods tend to return cluster numbers close to the maximal or minimal allowed value. Features that include the position of a pixel yield rather bad results. The suggested optimal number of clusters from the GAP criterion coincides best with our expectations. The corresponding reconstruction errors are also quite good, considering that the reconstruction error with the original 170 colours had an error of 46.96

In the following, we investigate how replacing the equidistant quantisation with our clustering approach influences the performance of this codec. Since k-means++ has yielded the most consistent results in the previous sections, we apply this strategy.

In our first experimental set-up, we use the same binary mask to compare the compression performance of equidistant and clustering-based quantisation: We quantise the original grey values with each approach, respectively, and apply the same coding techniques afterwards.

Figure 5 shows that for low compression ratios, equidistant and clustering-based quantisation lead to similar results with a slight benefit for clustering. However, for higher compression ratios, the equidistant quantisation yields clearly superior results. In order to understand this outcome, we analyse the dependencies of the reconstruction error and the ratio on the number of quantisation levels for both methods in Figs. 6 and 7. Compared to transform-based codecs, k-means does not quite reach the quality of JPEG. In contrast, the equidistant quantisation is on par with JPEG and it has even been shown to be able to beat JPEG2000 depending on compression ratio and image content (see [28]).

Figure 6 shows that the reconstruction error of clustering-based quantisation fluctuates much less depending on the number of quantised grey values than in the case of equidistant quantisation. However, it does not offer a distinct qualitative advantage. On the contrary, for very coarse quantisations, the equidistant quantisation can also outperform the results of clustering in isolated cases.

More importantly, the nonequidistant quantisation does not only change the reconstruction error, but also the entropy of the sequence of grey values that need to be stored. Figure 7 reveals that this yields a consistent increase in storage costs compared to an equidistant approach. Consequently,
clustering needs to use a coarser quantisation at the cost of reconstruction quality in order to reach the same compression ratio as an equidistant method. Overall, this leads to the superior performance of the equidistant quantisation.

Moreover, there is another factor that needs to be considered for our evaluation: State-of-the-art compression codecs do not only carefully select the location of known data, but also optimise the pixel values under the constraint of the quantisation [28, 31]. Data optimisation schemes, such as the method of Hoeltgen et al. [18], are unable to handle quantisations of the co-domain. Nevertheless, there is a clear mutual dependency between optimal data positions and values that should be respected in the quantisation process. Such a so-called quantisation-aware tonal optimisation efficiently corrects suboptimal data locations for the quantised colour values in a post-processing scheme and leads to a much more distinct advantage of equidistant quantisations (see Fig. 5). The clustering limits the ability of the tonal optimisation to adjust the behaviour of the inpainting algorithm locally. This is due to an inherent property of nonequidistant quantisations: By definition, the clustering leads to sparse regions in the grey value domain, where only few cluster centres are located, while other regions are densely populated. In isolated regions, quantisation-aware tonal optimisation can only apply large changes for grey values, while much smaller changes are possible in dense regions. This introduces a bias to the tonal optimisation that does not exist for equidistant quantisation: Here, each pixel value can be tuned by the same step size between quantisation levels and thereby allows to diverge more significantly from the original distribution of grey values.

Our evaluation shows that clustering on its own is not a suitable replacement for the simple equidistant quantisa-
tion in PDE-based compression. More complex quantisation techniques would require a more complex integration in the full compression pipeline, taking into account the balance of storage cost and reconstruction quality. While such an approach to nonequidistant quantisation would be promising, it is beyond the scope of our publication. Still, clustering could be a viable alternative due to its favourable run time. The \( k \)-means and hierarchical clustering have run times below one second for the presented test data. If the clustering is done on the mask without any positional information, then the run times are in the vicinity of 0.1 s. The evaluation of the quality criteria is more time-consuming though DB and CH are usually being fastest with run times of a several seconds and GAP often exceeding several minutes. The original tonal optimisation in R-EED [31] requires \( \approx 10 \) h on the same machine for full convergence. Fast quantisation-aware tonal optimisation [28] is considerably faster (8.2 s), but requires the pre-computation of so-called diffusion echos (4.07 m). All tests were done on a standard desktop PC with an Intel Core i7-6700CPU@3.4Ghz CPU.

5 Summary and conclusions

Clustering makes a lot of sense for pure quantisation of existing data. Our experiments show that we can reduce the number of colours used for PDE-based inpainting by as much as 80% without encountering any significant loss in the reconstruction quality. In addition, the presented strategies are fast and easy to implement. However, the cost of the associated data and the limitations imposed on tonal optimisation make this concept difficult to apply to compression if it is treated as an isolated component in the compression pipeline. Our findings show that a good quantisation with respect to the reconstruction error does not necessarily imply an efficient encoding of the data.

Therefore, successful data optimisation strategies in the context of image compression must include the encoding costs in their framework. It is not enough to consider the reconstruction error for the quantisation. Such potential new models are challenging in several ways. Not only are the encoding costs difficult to predict, but also their optimisation is likely to be costly. For practical applications it is however necessary to devise strategies that are very fast since state-of-the-art data compression codecs have run times in the millisecond range.

References

1. Belhachmi, Z., Bucur, D., Burgeth, B., Weickert, J.: How to choose interpolation data in images. SIAM J. Appl. Math. 70(1), 333–352 (2009)

5. Calinski, R.B., Harabasz, J.: A dendrite method for cluster analysis. Commun. Stat. 3, 1–27 (1974)

2. Bibiane-Schönlieb, C.: Partial Differential Equation Methods for Image Inpainting. Cambridge University Press, Cambridge (2015)

3. de Boor, C.: Good approximation by splines with variable knots II. In: Watson, G. (ed.) Conference on the Numerical Solution of Differential Equations, LNM, vol. 363, pp. 12–20. Springer (1974)

4. Bourguard, A., Unser, M.: Anisotropic interpolation of sparse generalized image samples. IEEE Trans. Image Process. 22(2), 459–472 (2013)
22. Lloyd, S.P.: Least squares quantization in pcm. IEEE Trans. Inf. Theory 28(2), 129–137 (1982)
23. Mahoney, M.: Adaptive weighing of context models for lossless data compression. In: Technical Report CS-2005-16, Florida Institute of Technology, Melbourne, Florida (2005)
24. Mainberger, M., Hoffmann, S., Weickert, J., Tang, C.H., Johannsen, D., Neumann, F., Doerr, B.: Optimising spatial and tonal data for homogeneous diffusion inpainting. In: Bruckstein, A.M., Haar ter Romeny, B.M., Bronstein, A.M., Bronstein, M.M. (eds.) Scale Space and Variational Methods in Computer Vision, LNCS, vol. 6667, pp. 26–37. Springer (2012)
25. Ochs, P., Chen, Y., Brox, T., Pock, T.: iPiano: Inertial proximal algorithm for non-convex optimization. SIAM J. Imag. Sci. 7(2), 1388–1419 (2014)
26. Ouni, T., Lassoued, A., Abid, M.: Lossless image compression using gradient based space filling curves (g-sfc). Signal Image and Video Process. 9(2), 277–293 (2015)
27. Peter, P., Hoffmann, S., Nedwed, F., Hoeltgen, L., Weickert, J.: From optimised inpainting with linear PDEs towards competitive image compression codecs. In: Briand, T., McCane, B., Rivers, M., Yu, X. (eds.) Advances in Image and Video Technology, LNCS, vol. 9431, pp. 63–74. Springer (2015)
28. Peter, P., Hoffmann, S., Nedwed, F., Hoeltgen, L., Weickert, J.: Evaluating the true potential of diffusion-based inpainting in a compression context. Signal Process. Image Commun. 46, 40–53 (2016)
29. Peter, P., Schmaltz, C., Mach, N., Mainberger, M., Weickert, J.: Beyond pure quality: progressive mode, region of interest coding and real time video decoding in PDE-based image compression. J. Visual Commun. Image Represent. 31, 253–265 (2015)
30. Rousseeuw, P.J.: Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Comput. Appl. Math. 20, 53–65 (1987)
31. Schmaltz, C., Peter, P., Mainberger, M., Ebel, F., Weickert, J., Bruhn, A.: Understanding, optimising, and extending data compression with anisotropic diffusion. Int. J. Comput. Vis. 108(3), 222–240 (2014)
32. Schmaltz, C., Weickert, J., Bruhn, A.: Beating the quality of JPEG 2000 with anisotropic diffusion. In: Pattern Recognition, LNCS, vol. 5748, pp. 452–461. Springer (2009)
33. Sundberg, R.: Maximum likelihood theory for incomplete data from an exponential family. Scand. J. Stat. 1(2), 49–58 (1974)
34. Sundberg, R.: An iterative method for solution of the likelihood equations for incomplete data from exponential families. Commun. Stat. Simul. Comput. 5(1), 55–64 (1976)
35. Tibshirani, R., Walther, G., Hastie, T.: Estimating the number of clusters in a data set via the gap statistic. J. R. Stat. Soc. B 63(2), 411–423 (2001)
36. Ward, J.H.: Hierarchical grouping to optimize an objective function. J. Am. Stat. Association 58, 236–244 (1963)
37. Zaremba, S.: Sur un problème mixte relatif à l’équation de Laplace. Bulletin de l’Académie des Sciences de Cracovie pp. 313–344 (1910)
38. Zeng, G., Ahmed, N.: A block coding technique for encoding sparse binary patterns. IEEE Trans. Acoust. Speech Signal Process. 37(5), 778–780 (1989)