

Fast Multilevel CVT-Based Adaptive Data Visualization Algorithm

M. Emelianenko*

Department of Mathematical Sciences, George Mason University, Fairfax, VA 22031, USA.

Received 25 September 2009; Accepted (in revised version) 6 January 2010

Available online 8 March 2010

Abstract. Efficient data visualization techniques are critical for many scientific applications. Centroidal Voronoi tessellation (CVT) based algorithms offer a convenient vehicle for performing image analysis, segmentation and compression while allowing to optimize retained image quality with respect to a given metric. In experimental science with data counts following Poisson distributions, several CVT-based data tessellation algorithms have been recently developed. Although they surpass their predecessors in robustness and quality of reconstructed data, time consumption remains to be an issue due to heavy utilization of the slowly converging Lloyd iteration. This paper discusses one possible approach to accelerating data visualization algorithms. It relies on a multidimensional generalization of the optimization based multilevel algorithm for the numerical computation of the CVTs introduced in [1], where a rigorous proof of its uniform convergence has been presented in 1-dimensional setting. The multidimensional implementation employs barycentric coordinate based interpolation and maximal independent set coarsening procedures. It is shown that when coupled with bin accretion algorithm accounting for the discrete nature of the data, the algorithm outperforms Lloyd-based schemes and preserves uniform convergence with respect to the problem size. Although numerical demonstrations provided are limited to spectroscopy data analysis, the method has a context-independent setup and can potentially deliver significant speedup to other scientific and engineering applications.

AMS subject classifications: 65D99, 65C20

Key words: Centroidal Voronoi tessellations, computational algorithms, Lloyd's method, acceleration schemes, multilevel method, binning, image analysis, visualization, signal-to-noise ratio.

1. Introduction

Centroidal Voronoi tessellations have diverse applications in many areas of science and engineering and the development of efficient algorithms for their construction is a key to their success in practice. Its use in imaging applications is the subject of many

*Corresponding author. *Email address:* memelian@gmu.edu (M. Emelianenko)

recent studies (e.g. [2–4] and references therein). This work will demonstrate how this concept can bring significant advantages to applications dealing with data visualization of extremely large data sets.

In particular, we focus on the problem of adaptively binning intensity images in physics applications, such as those coming from spectroscopy data with counting (Poisson) statistics. Hardness ratio maps, temperature maps and other types of images can be analyzed in a much similar manner. Spectroscopy has established itself as an unrivaled technique for a wide variety of scientific problems. However, experiments carried out using such techniques suffer from severe intensity limitations inherent to imperfect measurement instruments. Because of these limitations one usually needs to trade off signal-to-noise ratio (SNR) and instrumental resolution to raise statistical significance of measured data, which leads to the need to bin (or group) spatially close data for a better view of the image as a whole — a process referred to as *binning*. Adaptive techniques, able to relate bin sizes to local signal-to-noise ratios, are of extreme importance, since they prevent loss of resolution in high intensity areas. Even with adaptive binning, the size of the data to be processed is very large, with a typical modern day spectrometer delivering data sets of the order of 10^8 detector pixels. Hence high-speed and high-accuracy computational algorithms are crucial in order to make online data visualization and assessment possible.

This work builds upon several recently developed CVT-based methods for binning scattering data. Capellari and Copin [5] working on astrophysics imaging data were the first to develop a CVT-based adaptive binning method that achieves a homogeneous distribution of SNR across the image. Diehl and Statler [6] improved its performance by utilizing the concept of a weighted Voronoi diagram for added flexibility. Recently, Bustinduy et al. [7] proposed another adaptive algorithm that is able to preserve high-SNR features and avoid blurring common to the previous two approaches. While pursuing slightly different goals, all of the above works share one common drawback, which is high computational complexity associated with constructing optimal centroidal Voronoi tessellations of the given dataset. This fact serves as a motivation for the current study.

We suggest that a multidimensional extension of a multilevel scheme previously developed and extensively analyzed in 1-dimensional quantization context can significantly accelerate these and other data visualization techniques. The method was originally devised in one space dimension in [1], where it was proved to have uniform convergence with respect to the grid size for a large class of densities. Numerical studies of its behavior for some 2-dimensional domains with simplified geometries have been carried out in [1] and [8] with similar results, so the method is conjectured to have superior convergence properties regardless of dimension. Other possible acceleration techniques such as hybrid Lloyd-Newton and quasi-Newton algorithms, were considered in [8] and their applicability to imaging applications are subject to current explorations. In the version adapted for image analysis, the multilevel method is presently tested in the Capellari-Copin setting which aims at minimizing the spread of the signal-to-noise ratio around a target value $(S/N)_T$, but it can be very easily reformulated to fit other scenarios. More importantly, the multilevel algorithm description provided in current work is independent of the problem dimension or the features of the discrete dataset.

It has to be noted that an idea of using multiple grid levels to represent the scattering data has appeared in a previous work by Bustinduy et al [9] where they introduced a “multiresolution” adaptive binning algorithm that operates on several sequentially coarsened grids. The method serves a different purpose than that discussed here, and being restricted to rectangular grids suffers from very large variation of bin sizes (factors of 2^d , with d being the space dimension), which limits its applicability. We propose a completely different use of the multigrid context by suggesting a multilevel method that is able to:

- optimize the quality of the image in any given metric due to the flexibility and optimality of CVT-based grids;
- accelerate convergence comparing to other CVT-based schemes by replacing the costly Lloyd iteration with a more efficient alternative.

The paper is organized as follows. Section 2 introduces some necessary terminology and discusses computational features related to the Lloyd iteration employed by traditional adaptive binning algorithms. It is followed by Section 3 where we build a higher-dimensional extension of the multilevel algorithm. Application of the method to intensity image analysis, implications related to the discrete nature of the data and the description of Multilevel CVT-based Binning algorithm (MCVTBin) are given in Section 4. Results of numerical experiments and performance tests for the new method are summarized in Section 5.

2. Background and notations

A tessellation represents a mapping of N -dimensional vectors \mathbf{z}_i in the domain $\Omega \subset \mathbb{R}^N$ into a finite set of vectors $\{\mathbf{z}_i\}_{i=1}^k$. A Voronoi tessellation associates with each $\mathbf{Z} = \mathbf{z}_i$, also called a *generator*, a nearest neighbor region that is called a Voronoi region $\{V_i\}_{i=1}^k$. That is, for each i , V_i consists of all points in the domain Ω that are closer to $\mathbf{Z} = \mathbf{z}_i$ than to all the other generating points, and a Voronoi tessellation refers to the tessellation of a given domain by the Voronoi regions $\{V_i\}_{i=1}^k$ associated with a set of given generating points $\{\mathbf{z}_i\}_{i=1}^k \subset \Omega$. For a given density function ρ defined on Ω , we may define the centroids, or mass centers, of regions $\{V_i\}_{i=1}^k$ by

$$\mathbf{z}_i^* = \left(\int_{V_i} \mathbf{y} \rho(\mathbf{y}) d\mathbf{y} \right) \left(\int_{V_i} \rho(\mathbf{y}) d\mathbf{y} \right)^{-1}. \quad (2.1)$$

Then, an *optimal tessellation* may be constructed through a *centroidal Voronoi tessellation* (CVT) for which the generators of the Voronoi tessellation are the centroids of their respective Voronoi regions, in other words, $\mathbf{z}_i = \mathbf{z}_i^*$ for all i . An example of such a tessellation is given in Fig. 1(b), where it is compared to the regular Voronoi tessellation for the same choice of the density functional (a) and to a CVT with a different choice of density.

Given a set of points $\{\mathbf{z}_i\}_{i=1}^k$ and a tessellation $\{V_i\}_{i=1}^k$ of the domain, we may define

the *energy functional* or the *distortion value* for the pair $(\{\mathbf{z}_i\}_{i=1}^k, \{V_i\}_{i=1}^k)$ by:

$$\mathcal{G}(\{\mathbf{z}_i\}_{i=1}^k, \{V_i\}_{i=1}^k) = \sum_{i=1}^k \int_{V_i} \rho(\mathbf{y}) |\mathbf{y} - \mathbf{z}_i|^2 d\mathbf{y}. \quad (2.2)$$

The minimizer of \mathcal{G} , that is, the optimal tessellation, necessarily forms a CVT which illustrates the optimization property of the CVT [10].

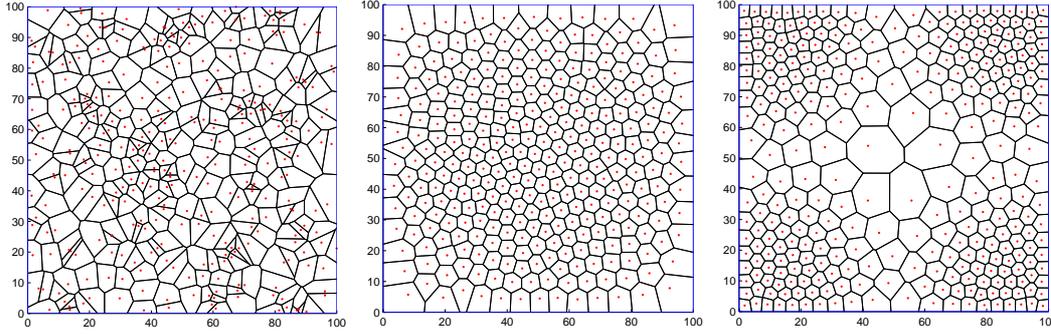


Figure 1: (a) Voronoi tessellation for $\rho = \exp(-(x-50)^2 + (y-50)^2)/1000$; (b) CVT for $\rho = \exp(-(x-50)^2 + (y-50)^2)/1000$; (c) CVT for $\rho = \exp-(x^2 + y^2)/500 + \exp-((x-100)^2 + y^2)/500 + \exp-(x^2 + (y-100)^2)/500 + \exp-((x-100)^2 + (y-100)^2)/500$.

In [11], Lloyd proposed the following iterative procedure for computing optimal tessellations: starting from an initial quantization (a Voronoi tessellation corresponding to an old set of generators), a new set of generators is defined by the mass centers of the Voronoi regions. This process is continued until certain stopping criterion is met. It is easy to see that the Lloyd algorithm is an energy descent iteration of the energy functional (2.2), which gives strong indications to its practical convergence.

Lloyd's algorithm and its variants have been proposed and studied in many contexts for different applications [2, 3, 12–18]. A particular extension using parallel and probabilistic sampling was given in [19] which allows efficient and mesh free implementation of the Lloyd's algorithm. Still, Lloyd algorithm is at best linearly convergent, besides it slows down as the number of generators gets large. In fact, it was shown in [20] that most smooth densities, convergence factor ρ satisfies

$$\rho \approx 1 - \frac{c}{k^2},$$

where k is the number of generators. We refer to [20] and [21] for some recent results and development of a rigorous convergence theory.

In [1], an energy-based multilevel acceleration scheme has been proposed that was shown to have a uniform convergence with respect to the problem size in 1-dimensional case. We are now ready to outline its higher-dimensional implementation and demonstrate its superior performance for some data visualization examples.

3. Optimization-based nonlinear multilevel algorithm

Since the original concept of centroidal Voronoi tessellations is related to the solution of a nonlinear optimization problem, and the monotone energy descent property is preserved by the Lloyd's iteration [10], it is natural to explore whether monotone energy reduction can be achieved in a multilevel procedure which would also improve the performance of the simple-minded iteration.

The problem of constructing a CVT is nonlinear in nature, hence standard linear multigrid theory cannot be directly applied. There are still several ways one could implement a nonlinear multilevel scheme in this context (see [8, 22, 23]). A Newton type acceleration method, studied earlier in [8], is based on some global linearization as the outer loop, coupled with other fast solvers in the inner loop. As an alternative, we studied an approach that overcomes the difficulties of the nonlinearity by essentially relying on the direct energy minimization without any type of global linearization.

Let us define the energy functional

$$\mathcal{H}(\{\mathbf{z}_i\}_{i=1}^k) = \mathcal{G}(\{\mathbf{z}_i\}_{i=1}^k, \{V_i\}_{i=1}^k), \quad (3.1)$$

where $\{V_i\}_{i=1}^k$ forms the Voronoi tessellation with generators $\{\mathbf{z}_i\}_{i=1}^k$. The CVTs and optimal quantizers are closely related to the problem of minimization of the functional \mathcal{H} . In fact, we may note that the vector of generators of a CVT forms a critical point of \mathcal{H} and vice versa [10]. That is, at a CVT (or optimal quantizer), we have $\nabla \mathcal{H} = 0$. This is one of the important characterizations of the CVTs which will be used in the later discussion.

In developing the multilevel quantization method, we followed the ideas presented in the literature on the extension of multigrid ideas to nonlinear optimization problems (see [24, 25] and the references cited therein).

3.1. Space decomposition

Since the energy functional is in general non-convex, it turns out to be very effective to relate our problem to an equivalent optimization problem through a technique that mimics the role of a dynamic nonlinear preconditioner. More precisely, denote $R = \text{diag}(R_1^{-1}, R_2^{-1}, \dots, R_k^{-1})$ a diagonal matrix whose diagonal entries $\{R_i = \int_{V_i} \rho(\mathbf{y}) d\mathbf{y}\}$ are the masses of the corresponding Voronoi cells. It is easy to deduce that $R\nabla \mathcal{H} = 0$ at a CVT. Hence we arrive at an equivalent formulation of the minimization problem: $\min \|R\nabla \mathcal{H}\|^2$, with respect to the standard Euclidean norm.

A key observation is that as R varies with respect to the generators, the above transformation or *dynamic preconditioning* keeps the modified objective functional convex in a suitably large neighborhood of the minimizer and therefore makes the new formulation more amenable to analysis than the original problem. Now, if we define the set of iteration points \mathbf{W} by

$$\mathbf{W} = \left\{ (w_i)_{i=1}^k \mid w_i \in \Omega, \forall 1 \leq i \leq k \right\},$$

we can design a new multilevel algorithm based on the following nonlinear optimization problem

$$\min_{\mathbf{Z} \in \mathbf{W}} \tilde{\mathcal{H}}(\mathbf{Z}), \text{ where } \tilde{\mathcal{H}}(\mathbf{Z} = \{\mathbf{z}_i\}_{i=0}^{k+1}) = \|\mathbf{R}\nabla \mathcal{H}(\{\mathbf{z}_i\}_{i=1}^k)\|^2. \quad (3.2)$$

The functional $\tilde{\mathcal{H}}$ may be regarded as a dynamically preconditioned energy. Viewing \mathbf{W} as a set of grid points in Ω , we may denote by $\mathcal{T} = \mathcal{T}_J$ a finite element mesh corresponding to \mathbf{W} , and consider a sequence of nested quasi-uniform finite element meshes $\mathcal{T}_1 \subset \mathcal{T}_2 \subset \dots \subset \mathcal{T}_J$, where \mathcal{T}_i consists of all finite element meshes $\{\tau_j^i\}_{j=1}^{n_i}$ with mesh parameter h_i , such that $\cup_{j=1}^{n_i} \tau_j^i = \Omega$. Corresponding to each finite element partition \mathcal{T}_i ($i = 1, \dots, J$), there is a finite element space \mathbf{W}_i defined by

$$\mathbf{W}_i = \left\{ v \in H^1(\Omega) \mid v|_{\tau} \in \mathcal{P}_1(\tau), \forall \tau \in \mathcal{T}_i \right\},$$

where H^1 is a Sobolev space with the usual norm $\|f\|_{H^1}^2$, and \mathcal{P}_1 denotes a space of piecewise linear functionals.

For each \mathbf{W}_i , there corresponds a nodal basis $\{\psi_j^i\}_{j=1}^{n_i}$, such that $\psi_j^i(x_k^i) = \delta_{jk}$, where δ_{jk} is the usual Kronecker Delta function and $\{x_k^i\}_{k=1}^{n_i}$ is the set of all nodes of the elements of \mathcal{T}_i . Define the corresponding one-dimensional subspaces $\mathbf{W}_{i,j} = \text{span}\{\psi_j^i\}$. Then the decomposition can be regarded as

$$\mathbf{W}_J = \sum_{i=1}^J \sum_{j=1}^{n_i} \mathbf{W}_{i,j} = \bigoplus_{i=1}^J \bar{\mathbf{W}}_i,$$

where

$$\bar{\mathbf{W}}_i = \mathbf{W}_i / \mathbf{W}_{i-1} \quad \text{for } i > 1 \quad \text{and} \quad \bar{\mathbf{W}}_1 = \mathbf{W}_1.$$

Now clearly for each $\psi_j^i \in \mathbf{W}_i$, we can find a vector $\bar{\psi}_j^i = \{\bar{\psi}_{jm}^i\} \in \mathbb{R}^{n_j}$, such that $\psi_j^i(x) = \sum_{m=1}^{n_j} \bar{\psi}_{jm}^i \psi_m^j(x)$, for $x \in \Omega$.

The set of basis functions

$$Q_i = [\bar{\psi}_1^i, \dots, \bar{\psi}_{n_i}^i]^T \in \mathbb{R}^{n_i \times k}$$

used at each iteration can be pre-generated using the recursive procedure:

$$Q_J = I_{k \times k} \quad \text{and} \quad Q_{J-s} = (\prod_{i=1}^s P_{J-i}) Q_J,$$

where P_i is the basis transformation from space \mathbf{W}_{i+1} to \mathbf{W}_i which plays a role of a restriction operator.

3.2. Description of the algorithm

Using the above notations, we design the Algorithm 3.1 which is a multilevel successive subspace correction algorithm. Each step of the procedure outlined below involves solving a system of nonlinear equations which plays the role of relaxation. We can use the Newton iteration to solve this nonlinear system. Solution at current iterate is updated after each nonlinear solve by the Gauss-Seidel type procedure, hence the resulting scheme is successive in nature. The ‘‘slash’’ cycle can be defined as described in Algorithm 3.1.

Algorithm 3.1. Successive correction $V(\nu_1, 0)$ scheme*Input:* k , number of generators; $u_1 = \{z_i\}_{i=1}^k$, the initial set of generators.*Output after n cycles:* $u_n = \{z_i\}_{i=1}^k$, the set of generators for CVT $\{V_i\}_{i=1}^k$.*Method:* For $n > 1$, given u_n , do1. For $i = 1 : J$

$$\bar{u}_{n+\frac{i-1}{J}} = u_{n+\frac{i-1}{J}}$$

For $l = 1 : \nu_1$

$$\bar{u}_{n+\frac{i-1}{J}} = \bar{u}_{n+\frac{i-1}{J}} + \alpha_{jl}^0 \bar{\psi}_j^i \in \bar{W}_i \text{ sequentially for } 1 \leq j \leq n_i,$$

such that

$$\mathcal{H}(\bar{u}_{n+\frac{i-1}{J}} + \alpha_{jl}^0 \bar{\psi}_j^i) = \min_{\alpha_{jl}} \mathcal{H}(\bar{u}_{n+\frac{i-1}{J}} + \alpha_{jl} \bar{\psi}_j^i),$$

endfor.

$$u_{n+\frac{i}{J}} = \bar{u}_{n+\frac{i-1}{J}} = u_{n+\frac{i-1}{J}} + e_n^i, \text{ where } e_n^i = \sum_{l=1}^{\nu_1} \sum_{j=1}^{n_i} \alpha_{jl}^0 \bar{\psi}_j^i$$

endfor.

2. On the coarsest level, $u_{n+1} \leftarrow \text{CoarseGridSolve}(u_{n+1})$.3. $n = n+1$

4. Repeat the procedure 1 to 3 until some stopping criterion is met.

Parameters ν_1, ν_2 denote the number of Gauss-Seidel iterations (smoothings) used at each level. Although it is enough to have $\nu_{1,2} = 1$ in theory, larger values need to be used in practice due to the numerical error in solving the nonlinear system. The values $\nu_{1,2} \leq 3$ usually suffice for the optimization to reach saturation. In the above description, no post-smoothings are used to make the analysis more transparent. A complete $V(\nu_1, \nu_2)$ cycle with ν_2 post-smoothings can be defined and analyzed in a similar fashion. The algorithm uses a procedure $\text{CoarseGridSolve}(\mathbf{Z})$, which, as the name indicates, refers to finding the solution at the coarsest level. In our implementation, this procedure consists of applying Lloyd method for a few steps or until saturation. In general, other efficient optimization methods, as well as Newton's method, can be used in order to quickly damp the error, since the number of unknowns on the coarsest grid remains relatively small.

The algorithm essentially only depends on the proper space decompositions and its relation to the set of generators, thus it is applicable in any dimension. Any particular implementation depends on the choice of nodes $\{x_k^i\}_{k=1}^{n_i}$ in the sequence of meshes $\mathcal{T}_1 \subset \mathcal{T}_2 \subset \dots \mathcal{T}_J$ and of the nodal basis functions $\{\psi_j^i\}_{j=1}^{n_i}$. These are the two major ingredients of a multilevel scheme that carry the most weight in defining its numerical properties. In what follows we will propose one possible coarsening and interpolation strategy and evaluate its numerical performance.

3.3. Coarsening procedure

Let \mathbf{W}_h be the vector space corresponding to the fine grid, and $\dim \mathbf{W}_h = n_h$. We would like to define \mathbf{W}_H with given $\dim \mathbf{W}_H = n_H$. Let us assume that we have in hand a full rank operator $P : \mathbb{R}^{n_H} \rightarrow \mathbf{W}_h$. Then a natural choice is $\mathbf{W}_H = \text{Range}(P)$.

Such a P can be made by fixing the supports of the basis functions, and then chopping these functions at the edges of the Voronoi diagram in order to get better properties of \mathbf{W}_H . The definition of the supports is based on the block tri-diagonal form of the graph Laplacian A . A variant of a breadth first search can be used to find a maximal independent set in the graph defined by A .

We start by using the boundary conditions to form the block structure of A , and assume that $\sum_{j=1}^n a_{ij} = 0$ except for the nodes “close” to the Dirichlet boundary. The algorithm then proceeds as follows:

Algorithm 3.2. Breadth First Search

1. Pick a set of starting nodes: $\mathcal{L}_0 = \{i \mid \sum_{j=1}^n a_{ij} \neq 0\}$. Set $k = 0$.
2. (a) Define $v \in \mathbb{R}^n$ such that $v_i = 1, i \in \mathcal{L}_j, ; v_i = 0, i \notin \mathcal{L}_j, j \leq k$.
(b) Define $\mathcal{L}_{k+1} = \{i \mid v_i = 0; (Av)_i \neq 0\}$. Set $k = k + 1$ and $v_i = v_i + (Av)_i$.
3. Until $\mathcal{L}_{k-1} = \emptyset$.

The set L_i contain indices of the nodes that have no edges connecting them to the points in L_{i-1} . This procedure defines “concentric layers” of nodes defined by the graph Laplacian of the matrix A . Such an implementation of the breadth first search algorithm requires $\mathcal{O}(n_j + |E|)$ operations and is optimal if A is sparse.

The next step consists in defining the coarse grid degrees of freedom, which is done in the way similar to the alternating construction used in 1d case. More precisely, we start by finding block tridiagonal form of A . We then remove all odd-numbered blocks and find block tridiagonal form of each connected component of each remaining even-numbered block. We repeat this process until only 1×1 and 2×2 blocks remain. The remaining evenly numbered “blocks” in the above procedure will be the coarse grid degrees of freedom.

This procedure picks coarse nodes from each of the layers L_i in such a way that coarse nodes are at least one node apart from each other in the graph defined by A , as demonstrated by Fig. 2.

3.4. Interpolation procedure

To make the presentation of our multilevel scheme complete, we need to define the interpolation operators that appeared in Algorithm 3.2. Namely, we need to discuss the choice of the restriction and prolongation operators. To define an interpolatory connection for each degree of freedom, which does not appear on the coarse grid, we have to uniquely determine the points on coarse grid which will be used to interpolate the value of the fine grid degree of freedom.

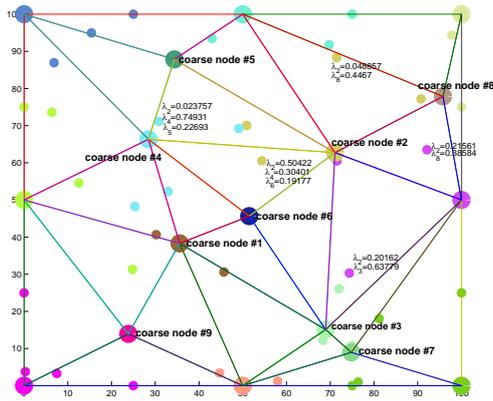


Figure 2: A visualization of the grid coarsening part of the algorithm.

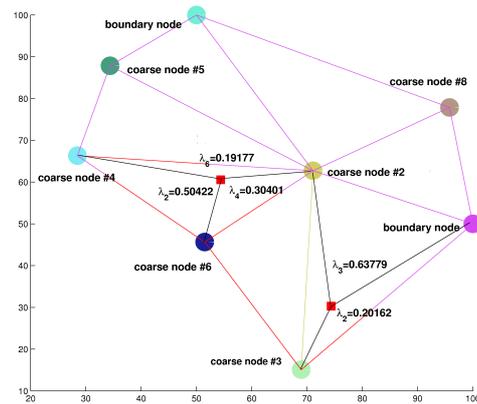


Figure 3: A schematic view of the interpolation based on barycentric coordinates.

In Section 3.1 we introduced the following space decomposition:

$$W_J = \sum_{i=1}^J \sum_{j=1}^{n_i} W_{i,j} = \bigoplus_{i=1}^J \bar{W}_i,$$

where $W_{i,j} = \text{span}\{\psi_j^i\}$ for the canonical nodal basis $\{\psi_j^i\}_{j=1}^{n_i}$. Based on this basis, we can define an interpolator the following way. At every level, define a triangulation based on coarse grid nodes. Then for each fine grid point identify the unique triangle it belongs to in this triangulation and set

$$\lambda_{i,j} = \{ \text{barycentric coordinate of a fine node } i \text{ with respect to coarse node } j \}.$$

Fig. 3 illustrates this definition.

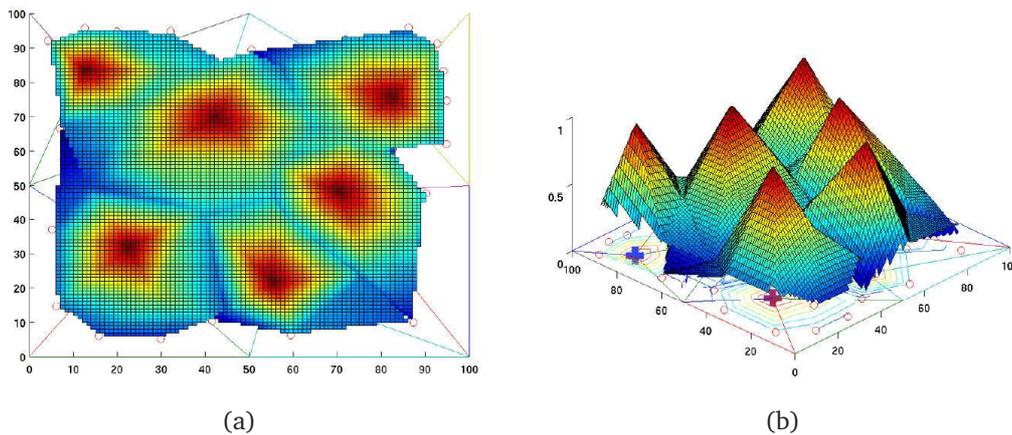


Figure 4: (a) Support of the basis functions on the coarse level; (b) Corresponding nodal basis.

The interpolation matrix $Q_i = [\bar{\psi}_1^i, \dots, \bar{\psi}_{n_i}^i]^T \in \mathbb{R}^{n_i \times k}$ is constructed as follows:

$$q_{j,i_c} = \lambda_{j,i_c}. \quad (3.3)$$

We further restrict the support of each basis function to the corresponding Voronoi cell as shown in Fig. 4.

A simplified version of this procedure has been originally tested on parallelogram shaped domains [1]. Here we presented the generalized version of this algorithm which we are now ready to adapt to the context of binning.

4. Accelerated Multilevel CVT-based Adaptive Binning (MCVTBin)

According to Gersho's conjecture [26], for bounded and strictly positive densities, the size of the Voronoi intervals are inversely proportional to the one-third power of the underlying density at the midpoints of the intervals. This conjecture can be used to partition the domain into equal SNR regions. In a 2-dimensional problem, a CVT with a density function $\rho = (S/N)^2$, generates a tessellation with bins that asymptotically enclose a constant mass (constant SNR). The obtained VT is strongly dependent on the discrete nature of input data, which can pose problems with degenerate cells and pinned pixels. A way out proposed in [5] is to use a special initial configuration of generators produced by the so-called **bin-accretion algorithm**, which alleviates these issues as long as the underlying density is additive, i.e.,

$$(S/N)_{1+2}^2 = (S/N)_1^2 + (S/N)_2^2.$$

The latter is true for all intensity maps subject to Poisson-distributed noise levels, which is the case under consideration. In order to adapt the multilevel scheme to the scope of discrete problems of this type, we adhere to the same strategy of producing a suitable starting configuration, as outlined below.

Bin accretion starts by choosing the highest S/N pixel in the image. We then find centroid of the current bin and select the unbinned pixel closest to the centroid. We add the candidate to a chosen bin if:

- (a) the new pixel is adjacent to the current bin;
- (b) the morphology of the bin would remain below some threshold;
- (c) the S/N for this bin would get closer to $(S/N)_T$.

The process is repeated until no new pixel can be added. Let k_i denote the number of pixels in V_i . If the bin's S/N computed as

$$(S/N)_{V_i} = \frac{\sum_{i=1}^{k_i} S_i}{\sqrt{\sum_{i=1}^{k_i} N_i^2}},$$

is higher than some fraction of the target signal-to-noise ratio, e.g. $0.3(S/N)_T$, we mark all the pixels in the bin as successfully binned, otherwise as unsuccessfully binned. We find mass centroid of all pixels already binned and start a new bin from an unbinned pixel closest to this centroid, then repeat until all pixels have been binned. Then we compute

centroid of each successful bin and reassign the unsuccessfully binned pixels to the closest of these centroids. It remains to recompute centroids of each bin obtained above and use them as input for the multilevel $V(v_1, 0)$ -cycle defined in Algorithm 3.1. This strategy serves as a first step of the main algorithm proposed in this study — Multilevel CVT-based Binning (MCVTBin) algorithm — which is outlined below.

Algorithm 4.1. MCVTBin algorithm.

1. Run bin accretion to generate initial configuration $\mathbf{Z} = \{z_i\}_{i=1}^k$.
2. Perform a multilevel cycle $V(v_1, 0)$ described in Algorithm 3.1:

For $i = 1 : J$

 - (a) Use Algorithm 3.2 to define the partition \mathcal{T}_i and coarse degrees of freedom.
 - (b) Use formula (3.3) to compute nodal basis $\{\psi_j^i\}_{j=1}^{n_i}$.
 - (c) Carry out v_1 smoothings $\tilde{\mathcal{H}} \rightarrow \min$, find weights α_{j_l} and update the solution. Use `CoarseGridSolve` to obtain nearly exact solution on the coarsest level.
3. Repeat Step 2 until one of the convergence criteria is met.

Convergence criterion adopted in this implementation of the method is a superposition of the following conditions: (a) $\|\mathbf{Z}_n - \mathbf{Z}_{n-1}\| < \epsilon_1$, (b) $\|\tilde{\mathcal{H}}(\mathbf{Z}_n) - \tilde{\mathcal{H}}_{ex}\|/\tilde{\mathcal{H}}_0 < \epsilon_2$ and (c) $n > N_{iter}$, where N_{iter} defines the maximum number of multilevel cycles. Exact numerical values of the energy $\tilde{\mathcal{H}}_{ex}$ are not available in practice, but are used here to assess convergence and are obtained by running Lloyd iteration till saturation (a long time limit).

5. Numerical results

Below we analyze the performance of the MCVTBin algorithm using a set of two large-scale two-dimensional binning examples. The first one is a sample Integral-Field Spectroscopy (IFS) image provided by M. Cappellari (Oxford, UK), from the public repository at <http://www-astro.physics.ox.ac.uk/mxc/idl/#binning>. The data comes in a form of 3107 pixels, each with associated signal and noise values. Figs. 5(a,b) give a visualization of the original image and the result of the bin-accretion stage. The colormap is given by the SNR (signal-to-noise ratio), rescaled in the coarsened images to fit the range of remaining intensities for enhanced grid shape visualization. In Figs. 5(c,d) we compare the final binned image after 12 steps of the Lloyd algorithm (our implementation of the method of [5]), when the stopping criterion (a) was met, with that after 4 MCVTBin iterations, at which point criterion (b) was satisfied. The differences between the last two images are barely noticeable, which is to be expected due to the fact that both methods reach the same energy level at this point. It is clear that it takes MCVTBin less iterations than Lloyd to reach same accuracy level. The evolution of the energy and convergence estimates are given in Figs. 6 and 7. Convergence factor after n multilevel cycles is computed as $\rho \approx (\|e_n\|/\|e_0\|)^{1/n}$, with the error $e_n = \tilde{\mathcal{H}}(\mathbf{Z}_n) - \tilde{\mathcal{H}}_{ex}$ computed with respect to $\tilde{\mathcal{H}}_{ex}$ estimate obtained as a limit of Lloyd iterations with the same number of generators.

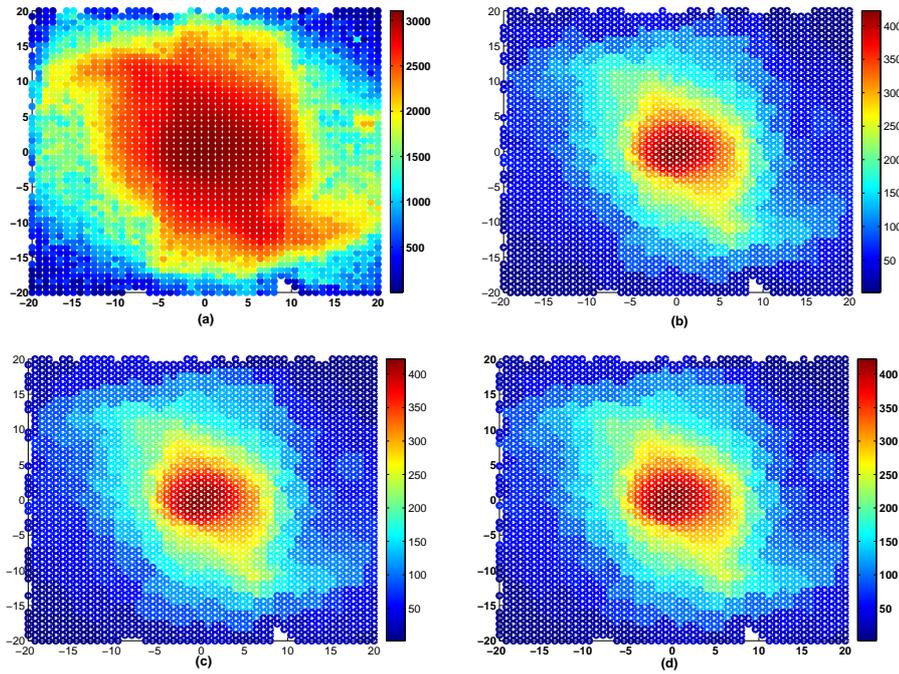


Figure 5: (a) Original IFS image from SAURON data. (b) The image after bin accretion. (c) Result of Lloyd algorithm run until saturation. (d) Result of MCVTBin algorithm after 4 steps.

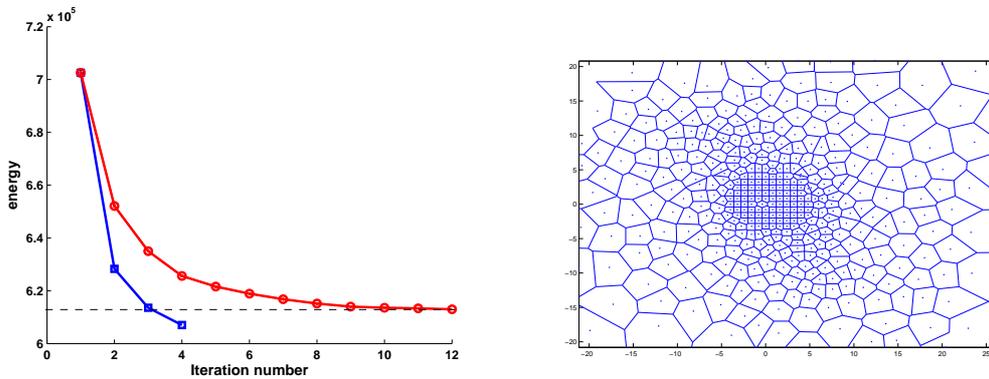


Figure 6: (a) Convergence of energies for Lloyd algorithm (crosses) vs. MCVTBin (squares), for the IFS example. (b) Final Voronoi tessellation provided by the multilevel solution for the same data.

Here we use the fact that since Lloyd method is globally convergent in the weak energy norm [1, 21], running Lloyd till stagnation yields a good approximation to the exact energy value. Reaching sufficiently low energy value is enough to ensure convergence of the algorithm in many practical applications, including the binning context considered here.

MCVTBin does require more computational time per cycle, but this fact is offset by the evidence of uniform convergence provided in Fig. 7. Computational savings are associ-

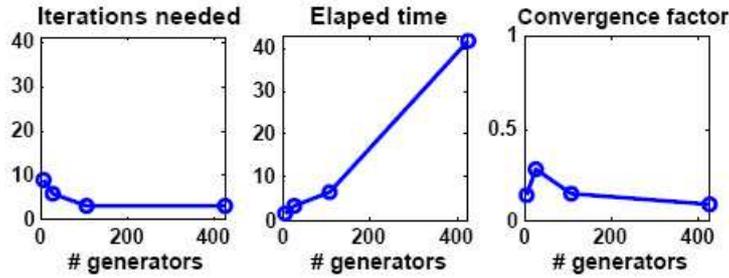


Figure 7: Convergence of the MCVTBin scheme for the IFS image: (a) the number of iterations required to reach the accuracy of $\epsilon = 10^{-6}$ does not grow, (b) computational time required to reach this accuracy, (c) convergence factor, computed per multilevel cycle.

ated with the fact that convergence of MCVTBin does not slow down with the number of generators in contrast with the deterioration of the Lloyd method's convergence. Fig. 6(b) shows the Voronoi tessellation obtained by the multilevel algorithm after 4 iterations.

Although the computational savings that we obtained in this example are not dramatic due to the size limitations of the original dataset, by the nature of the algorithm the advantage of the MCVTBin method becomes much more visible as the size of the problem increases. The next example provides some further evidence of this fact.

This time the set of tests comes from the analysis of the collective motions of magnetic moments within a slice of Cobalt. The data was obtained from the MAPS spectrometer at The Rutherford Appleton Laboratory (UK), courtesy of I. Bustinduy. It contains around 25k pixels with signal and noise values, given in the reciprocal space units, where $u_1 = [0.5Q_h, 0, 0, 0]$ and $u_2 = [0.5Q_h, -Q_k, 0, 0]$. Following standard notation, the indices h, k, l specify different crystal planes within the anisotropic structure of a crystalline solid, and the variables Q_h, Q_k, Q_l are the corresponding changes of momentum in 3D space. Fig. 8 gives the original image and the result after the bin-accretion stage of the algorithm. Fig. 9 provides a visualization of the image at different coarsening levels of the multilevel algorithm. The original colormap represents the intensity given as a SNR (neutron counts), while the colormap in coarsened images has been rescaled to precisely match the full spectrum of representative intensities.

The results of the benchmarking tests for this data are summarized in Fig. 10. As before, convergence factor is measured as the energy error reduction per cycle. MCVTBin produces same quality of binned image after a much smaller number of iterations (cf. [7, 9]) and, most importantly, preserves convergence factors when the number of pixels is increased. This means that the computational savings will increase even more as the size of the data grows.

From the above analysis we also conclude that computational time required by the MCVTBin algorithm scales linearly with k (number of generators, specifying the size of the problem) for both examples. The method maintains uniform convergence with respect to the problem size similar to its 1-dimensional prototype developed in [1], which gives a

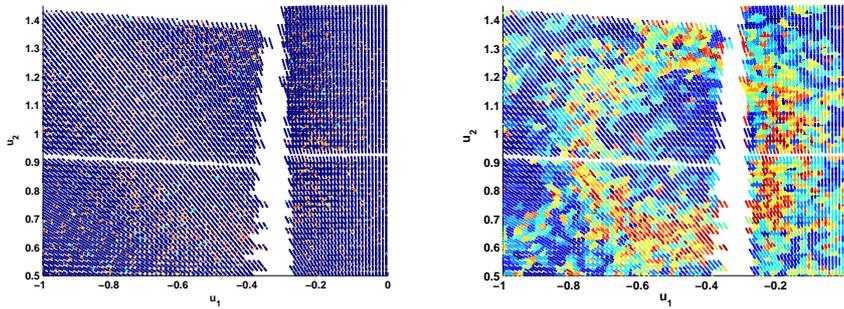


Figure 8: Collective motions of magnetic moments within cobalt, where the spin dispersion volume is sliced in the projection space defined by $u_1 = [0.5Q_h, 0, 0, 0]$ and $u_2 = [0.5Q_h, -Q_k, 0, 0]$: (left) Original experimental image of a magnetic cobalt system (25k pixels). (right) The image after bin accretion (2136 pixels). Colormap represents neutron counts.

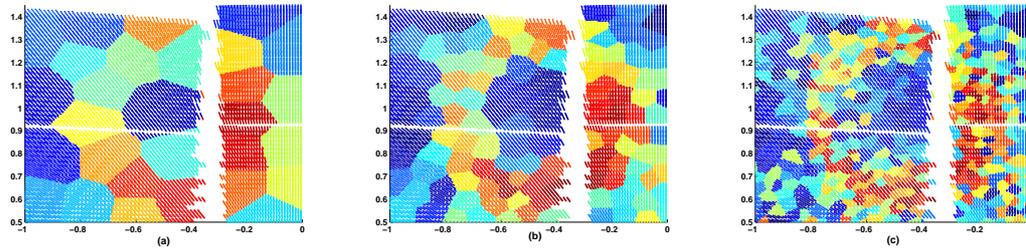


Figure 9: Comparison of grids at different coarse levels of the MCVTBin algorithm applied to the Cobaltslice data from MAPS: (a) The mesh at the lowest level of the multilevel algorithm (29 coarse nodes of freedom), (b) The mesh at level 2 (122 degrees of freedom), (c) The mesh at level 1 (512 degrees of freedom).

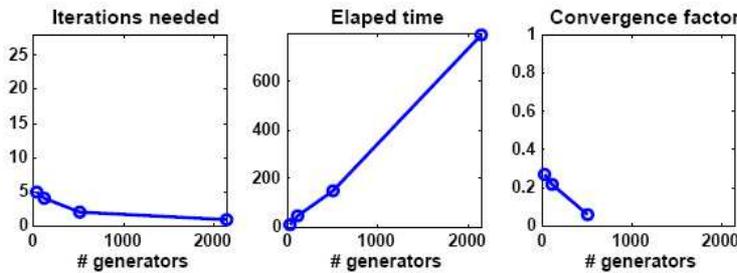


Figure 10: Convergence of the MCVTBin scheme for the Cobalt image: (a) the number of iterations required to reach the accuracy of $\epsilon = 10^{-6}$ does not grow, (b) computational time required to reach this accuracy, (c) convergence factor, computed per multilevel cycle. Since the method converges in 1 step when $k = 2136$, there is no convergence factor estimate available for this point in the rightmost graph.

significant advantage compared to other iterative methods. Calculations were carried out on a 3GHz machine running a MATLAB implementation of the algorithm, which can be made significantly faster when coded in higher level programming language.

6. Conclusion and future work

We have shown that the performance of CVT-based binning algorithms can be successfully optimized by means of multilevel accelerators such as the MCVTBin scheme introduced in this work. The extension of the multilevel quantization algorithm introduced in [1] to the two-dimensional case was made possible through barycentric coordinate based interpolation and maximal independent set coarsening procedures. Numerical evidence has been presented confirming the conjectured uniform convergence of the proposed scheme with respect to the problem size for two-dimensional examples with arbitrary densities, guaranteeing a significant speedup comparing with traditional methods. Care must be taken to produce an initial configuration allowing for guaranteed descent of the method into the area of fast convergence, which for example can be done through the widely known bin accretion technique. Note, however, that the bin accretion technique is important only when there is a danger of having on average very few data points inside each bin. This step can be completely avoided in case when the relative size of each Voronoi cell is large (e.g. ≈ 100 pixels per cell), with no toll on the convergence of the MCVTBin algorithm. The comprehensive analysis of convergence estimates for both regimes addressing the needs of different applications is currently underway.

In addition, it might be possible to extend the method to non-Poissonian noise distributions and try choosing different interpolation and coarsening rules, which can potentially have big impact on the algorithm performance. In particular, such modifications can have a positive or negative effect on the size of the convergence region - a question that can be rigorously investigated via employing ideas from the algebraic multigrid theory. Rigorous convergence analysis and application of multilevel acceleration techniques to other image and signal analysis problems is a focus of a separate study and will be discussed in details in future publications.

Acknowledgments The author wishes to thank anonymous referees for providing invaluable feedback that greatly improved the quality of the manuscript. The author is also indebted to Prof. Ludmil Zikatanov and Prof. Qiang Du of Penn State University for many fruitful discussions and help with the development of the original version of the multilevel algorithm, Zichao Di for assistance with running the code and Ibon Bustinduy for useful insights and for providing spectroscopy data that was used to benchmark the algorithm. The author was partially supported by the grants DMS 0405343 and DMR 0520425.

References

- [1] Q. Du and M. Emelianenko. Uniform convergence of a nonlinear energy-based multilevel quantization scheme via centroidal voronoi tessellations. *SIAM J. Numer. Anal.*, 46:1483–1502, 2008.
- [2] C. Wager, B. Coull, and N. Lange. Modelling spatial intensity for replicated inhomogeneous point patterns in brain imaging. *Journal of Royal Statistical Society B*, 66:429–446, 2004.

- [3] Q. Du, M. Emelianenko, Lee H.-C., and X. Wang. Ideal point distributions, best mode selections and optimal spatial partitions via centroidal voronoi tessellations. *Proc. 2nd Intl. Symp. Voronoi Diagrams in Sciences and Engineering (VD2005)*, pages 325–333, 2005.
- [4] Q. Du, M. Gunzburger, L. Ju, and X. Wang. Centroidal voronoi tessellation algorithms for image compression, segmentation, and multichannel restoration. *J. Math. Imaging Vis.*, 24(2):177–194, 2006.
- [5] M. Cappellari and Y. Copin. Adaptive spatial binning of integral-field spectroscopic data using voronoi tessellations. *Mon. Not. R. Astron. Soc.*, 342:345–354, 2003.
- [6] S. Diehl and T.S. Statler. Adaptive binning of x-ray data with weighted voronoi tessellations. *Mon. Not. R. Astron. Soc.*, 368:497–510, 2006.
- [7] I. Bustinduy, F.J. Bermejo, T.G. Perring, and G. Bordel. Experimental neutron spectroscopy data visualization: Adaptive tessellation algorithm. *Review of Scientific Instruments*, 78:234509, 2007.
- [8] Q. Du and M. Emelianenko. Acceleration schemes for computing the centroidal voronoi tessellations. *Numer. Linear Algebra Appl.*, 13:173–192, 2006.
- [9] I. Bustinduy, F.J. Bermejo, T.G. Perring, and G. Bordel. A multiresolution data visualization tool for applications in neutron time-of-flight spectroscopy. *Nuclear Inst. and Methods in Physics Research, A.*, 546:498–508, 2005.
- [10] Q. Du, V. Faber, and M. Gunzburger. Centroidal voronoi tessellations: applications and algorithms. *SIAM Review*, 41:637–676, 1999.
- [11] S. Lloyd. Least square quantization in pcm. *IEEE Transactions on Information Theory*, 28:129–137, 1982.
- [12] R. Gray and D. Neuhoff. Quantization. *IEEE Transactions on Information Theory*, 44:2325–2383, 1998.
- [13] L. Ju, M. Gunzburger, and Q. Du. Meshfree, probabilistic determination of points, support spheres, and connectivities for meshless computing. *Computational Methods in Applied Mechanics and Engineering*, 191:1349–1366, 2002.
- [14] Y. Linde, A. Buzo, and R. Gray. An algorithms for vector quantizer design. *IEEE Transactions on Communications*, 28:84–95, 1980.
- [15] J. MacQueen. Some methods for classification and analysis of multivariate observations. *Proc. 5th Berkeley Symp. Math. Stat. Prob., I, Edited by L. LeCam & J. Neyman*, pages 281–297, 1967.
- [16] A. Trushkin. On the design of an optimal quantizer. *IEEE Transactions on Information Theory*, 39:1180–1194, 1993.
- [17] Q. Du and M. Gunzburger. Grid generation and optimization based on centroidal voronoi tessellations. *Applied Mathematics and Computation*, 133:591–607, 2002.
- [18] J. Cortes, S. Martinez, T. Karata, and F. Bullo. Coverage control for mobile sensing networks. *IEEE Transactions on Robotics and Automation*, 20:243–255, 2004.
- [19] L. Ju, Q. Du, and M. Gunzburger. Probabilistic methods for centroidal voronoi tessellations and their parallel implementations. *Parallel Computing*, 28:1477–1500, 2002.
- [20] Q. Du, M. Emelianenko, and L. Ju. Convergence properties of the lloyd algorithm for computing the centroidal voronoi tessellations. *SIAM J. Numer. Anal.*, 44:102–119, 2006.
- [21] M. Emelianenko, L. Ju, and A. Rand. Weak global convergence of the lloyd method for computing centroidal voronoi tessellations in r^d . *SIAM J. Numer. Anal.*, 46:1423–1441, 2008.
- [22] Koren Y., Yavneh I., and A. Spira. A multigrid approach to the scalar quantization problem. *IEEE Transactions on Information Theory*, 51(8):2993–2998, 2005.
- [23] Koren Y. and Yavneh I. Adaptive multiscale redistribution for vector quantization. *SIAM Journal on Scientific Computing*, 27(5):1573–1593, 2006.

- [24] Q. Chang and Z. Huang. Efficient algebraic multigrid algorithms and their convergence. *SIAM Journal on Scientific Computing*, 24:597–618, 2002.
- [25] X. Tai and J. Xu. Global convergence of subspace correction methods for some convex optimization problems. *Mathematics of Computation*, 71:105–124, 2002.
- [26] A. Gersho. Asymptotically optimal block quantization. *IEEE Transactions on Information Theory*, 25:373–380, 1979.