Nonlinear and Non-Separable Multiscale Representations Based on Lipschitz Perturbation

Cédric Gérot [a] Basarab Mateï [b] Sylvain Meignen [c]

[a] *GIPSA Laboratory, University of Grenoble, France*
[b] *LAGA Laboratory, Paris XIII University, France*
[c] *LJK Laboratory, University of Grenoble, France*

Received *****; accepted after revision +++++

Presented by ££££££

Abstract

In this paper, we present a new formalism for nonlinear and non-separable multiscale representations. The new formalism we propose brings about similarities between existing nonlinear multiscale representations and also allows us to alleviate the classical hypotheses made to prove the convergence of the multiscale representations. *To cite this article: C. Gérot, B.Mateï and S. Meignen, C. R. Acad. Sci. Paris, Ser. I (2010).*

Résumé

Un nouveau formalisme pour les représentations multi-échelles non-linéaires et non-séparables. *Pour citer cet article :C. Gérot, B.Mateï and S. Meignen, C. R. Acad. Sci. Paris, Ser. I (2010).*

Version française abrégée

1. Introduction

Nonlinear multiscale representations are naturally defined using nonlinear prediction operators. In applications, it may be of interest to define multiscale representations that are not based on a dyadic grid. Several examples exist in image processing where the use of representations built using non-dyadic grids significantly improves the compression performance [2]. We define *Lipschitz-Linear* prediction operators in that context, and we give several examples of such operators, namely the recently introduced PPH

Email addresses: `cedric.gerot@gipsa-lab.inpg.fr` (Cédric Gérot), `matei@math.univ-paris13.fr` (Basarab Mateï), `sylvain.meignen@imag.fr` (Sylvain Meignen).

scheme. We give convergence and stability results of the nonlinear multiscale representations based on *Lipschitz-Linear* prediction operators both in L^p and Besov spaces. These results are part of a much deeper study available in [4]. A new aspect is introduced, namely the notion of prediction operators compatible with a set of differences. The results on the convergence and the stability of the corresponding multi-scale representations are identical to those obtained for representations based on *Lipschitz-Linear* prediction operators. For applications, the interesting aspect of this notion of compatibility is that it enables to reduce the complexity of the study of the joint spectral radius. We conclude the paper showing the convergence of some nonlinear multi-scale representations associated with *Lipschitz-Linear* prediction operators, namely the PPH scheme.
Before we start, we need to introduce some standard multi-index notations. For example, for $\alpha = (\alpha_1, \cdots, \alpha_d) \in \mathbb{N}_0^d$ we write $|\alpha| = \sum_{i=1}^d \alpha_i$ and for $x \in \mathbb{R}^d$ we write $x^\alpha = x_1^{\alpha_1} \cdots x_d^{\alpha_d}$, monomial with degree $|\alpha|$. By $(e_1, \cdots, e_d)$ we denote the canonical basis on $\mathbb{Z}^d$. The finite number of all monomials x^α with degree N is denoted by r_N^d. We then introduce $\prod_N$ the space of polynomials of degree N generated by $\{x^\alpha = \prod_{i=1}^d x_i^{\alpha_i}, |\alpha| \leq N\}$. In what follows, we will write $\deg(p)$ for the degree of any polynomial p. With that in mind, we denote, for any multi-index α and any sequence $(v_k)_{k\in\mathbb{Z}^d}$: $\Delta^\alpha v_k = \Delta_{e_1}^{\alpha_1} \cdots \Delta_{e_d}^{\alpha_d} v_k$ where $\Delta_{e_d}^{\alpha_d} v_k$ is defined recursively by: $\Delta_{e_d}^{\alpha_d} v_k = \Delta_{e_d}^{\alpha_d - 1} v_{k+e_d} - \Delta_{e_d}^{\alpha_d - 1} v_k$. For a given multi-index α, we will say that Δ^α is a difference of order $|\alpha|$. For any $N \in \mathbb{N}$, we will denote $\Delta^N v_k = \{\Delta^\alpha v_k, |\alpha| = N\}$.

2. Multiscale Representations

We assume that we are given the data $(v_k^j)_{k\in\mathbb{Z}^d}$, $j \geq 0$, associated to the locations $\Gamma^j = \{M^{-j}k, k \in \mathbb{Z}^d\}$, where M is a dilation matrix, (i.e. an invertible matrix in $\mathbb{Z}^d \times \mathbb{Z}^d$ satisfying $\lim\limits_{n\to+\infty} M^{-n} = 0$). We also consider that there exists a prediction operator S that computes $\hat{v}^j = Sv^{j-1}$, an approximation of v^j. Then, we define the prediction error as $e^j = v^j - \hat{v}^j$. The information contained in v^j is completely equivalent to (v^{j-1}, e^j). By iterating this procedure from the initial data v^J, we obtain its *nonlinear multiscale representation* $\mathcal{M}v^J = (v^0, e^1, \cdots, e^J)$ [3]. Conversely, assume that the sequence $(v^0, (e^j)_{j\geq 0})$ is given, we are interested in studying the convergence of the following nonlinear iteration:

$$v^j = Sv^{j-1} + e^j, \tag{2.1}$$

to a limit function v, which is defined as the limit (when it exists) of $v_j(x) = \sum\limits_{k\in\mathbb{Z}^d} v_k^j \varphi_{j,k}(x)$, where $\varphi_{j,k}(x)$ denotes $\varphi(M^j x - k)$ and φ is some compactly supported function satisfying the scaling equation:

$$\varphi(x) = \sum_{n\in\mathbb{Z}^d} g_n \varphi(Mx - n) \text{ with } \sum_n g_n = m := |\det(M)|. \tag{2.2}$$

When the sequence of functions $(v_j)_{j\geq 0}$ is convergent to some limit function in some functional space, by abusing a little bit terminology, we say that the multiscale representations $(v^0, (e^j)_{j\geq 0})$ is convergent in that space. The scaling equation (2.2) gives rise to the definition of a local and linear prediction operator S_l as follows:

$$S_l v_k = \sum_{l\in\mathbb{Z}^d} g_{k-Ml} v_l \tag{2.3}$$

3. *Lipschitz-Linear* Prediction Operators

In the following, a nonlinear prediction operator is a map $v \in \ell^\infty(\mathbb{Z}^d) \mapsto Sv \in \ell^\infty(\mathbb{Z}^d)$. In this paper, we study a particular type of nonlinear prediction operators defined by the sum of the linear prediction operator S_l and of a perturbation term which we define hereafter. The linear prediction operator shall satisfy the polynomial reproduction property which we now recall:

Definition 3.1 *We say that a prediction operator S reproduces polynomials of degree N if for $u_k = p(k)$ for any $p \in \prod_N$, we have*

$$Su_k = p(M^{-1}k) + q(k)$$

where q is a polynomial such that $\deg(q) < \deg(p)$. When $q = 0$, we say that the prediction operator exactly reproduces polynomials.

With this in mind, we introduce the definition of a *Lipschitz-Linear* prediction operator:

Definition 3.2 *A prediction operator S is Lipschitz-Linear of order $N+1$, if it is the sum of the linear prediction operator S_l, assuming to reproduce polynomials of degree N, and of perturbation terms defined using Lipschitz functions Φ_i $i = 0, \cdots, m-1$, such that $\Phi_i(0) = 0$, as follows:*

$$Sv_{Mk+i} = S_l v_{Mk+i} + \Phi_i(\Delta^{N+1} v_{k+p_1}, \cdots, \Delta^{N+1} v_{k+p_q}) \quad \forall i \in coset(M)$$

where $\{p_1, \cdots, p_q\}$ is a fixed set and where M is the dilation matrix associated to the definition of S_l.

From the above definition, we remark that when S_l reproduces polynomials of degree N so does S.

4. One-Dimensional *Lipschitz-Linear* Prediction Operators

We start by considering the one-dimensional case with $M = 2$. Given a set of embedded grids $\Gamma^j = \left\{2^{-j}k,\ k \in \mathbb{Z}\right\}$ we consider discrete values v_k^j defined on each vertex of these grids. These quantities shall represent a certain function v at level j. As an illustration, we show that the PPH scheme is an example of *Lipschitz-Linear* prediction operator ([1]). The PPH scheme is defined by: $\hat{v}_{2k+1}^j = \frac{v_{k+1}^{j-1} + v_k^{j-1}}{2} - \frac{1}{8} H(\Delta^2 v_{k-1}^{j-1}, \Delta^2 v_k^{j-1})$ and $\hat{v}_{2k}^j = v_k^{j-1}$ where $H(x, y) := \frac{xy}{x+y}(\text{sign}(xy) + 1)$. Since H satisfies $|H(x, y) - H(x', y')| \leq 2 \max\{|x - x'|, |y - y'|\}$, it is a Lipschitz function and since $|H(x, y)| \leq |\max(x, y)|$ it is bounded. Finally, the linear scheme $\frac{v_{k+1}^{j-1} + v_k^{j-1}}{2}$ reproduces polynomials of degree 1, therefore the PPH-scheme is a *Lipschitz-Linear* prediction operator of order 2.

5. Multi-Dimensional *Lipschitz-Linear* Prediction Operators on Non-Dyadic Grids

The motivation to consider non-dyadic grids are, for instance, better image compression results (see [2]). Having defined the grid $\Gamma^j = \left\{M^{-j}k, \quad k \in \mathbb{Z}^d\right\}$ using a dilation matrix M, we consider discrete quantities v_k^j defined on each of these grids. They shall represent a certain approximation of a function v at level j. As an illustration, we consider the bidimensional PPH-scheme defined by where the grids Γ^j are associated with the quincunx matrix M. We consider the following extension of the PPH scheme to the bidimensional case: $\hat{v}_{Mk+e_1}^j = \frac{v_k^{j-1} + v_{k+Me_1}^{j-1}}{2} - \frac{1}{8} H(\Delta_{Me_1}^2 v_k^{j-1}, \Delta_{Me_1}^2 v_{k-Me_1}^{j-1})$ and $\hat{v}_{Mk}^j = v_k^{j-1}$. Note that the linear part of the prediction operator is obtained by considering the affine interpolation polynomial at $v_k^{j-1}, v_{k+e_1}^{j-1}$ and $v_{k+e_1+e_2}^{j-1}$ and thus reproduces polynomials of degree 1. Since the perturbation is associated to a bounded Lipschitz function of the differences of order 2, this multi-dimensional prediction operator is *Lipschitz-Linear* of order 2.

6. Convergence Theorems

The convergence theorems are obtained by studying the difference operators associated to *Lipschitz-Linear* prediction operators. The existence of such difference operators is ensured by the following theorem:

Theorem 6.1 *Let S be a Lipschitz-Linear prediction operator of order $N+1$ then there exist multi-dimensional local operators $S^{(k)}$, $k \leq N+1$ such that $\Delta^k Sv = S^{(k)}\Delta^k v$*

To study the convergence of the iteration (2.1), we introduce the definition of the joint spectral radius for difference operators:

Definition 6.2 *Let us consider a Lipschitz-Linear prediction operator S of order $N+1$. The joint spectral radius in $(\ell^p(\mathbb{Z}^d))^{r_k^d}$ of $S^{(k)}$, for $k \leq N+1$ is given by*

$$\rho_p(S^{(k)}) := \inf_{j\geq 0} \|(S^{(k)})^j\|^{1/j}_{(\ell^p(\mathbb{Z}^d))^{r_k^d}\to(\ell^p(\mathbb{Z}^d))^{r_k^d}} = \inf_{j>0}\{\rho, \|\Delta^k S^j v\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}} \leq \rho^j \|\Delta^k v\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}}\}. \quad (6.1)$$

We then have the following

Theorem 6.3 *Let S be a Lipschitz-Linear prediction operator of order $N+1$. Assume that $\rho_p(S^{(k)}) < m^{1/p}$, for some $k \leq N+1$ and that $\|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{j>0} m^{-j/p}\|e^j\|_{\ell^p(\mathbb{Z}^d)} < \infty$. Then, the limit function v belongs to $L^p(\mathbb{R}^d)$ and*

$$\|v\|_{L^p(\mathbb{R}^d)} \leq C\left(\|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{j>0} m^{-j/p}\|e^j\|_{\ell^p(\mathbb{Z}^d)}\right) \text{ for some } C>0 \quad (6.2)$$

Remark 1 Usually the convergence in L^p is associated to the condition $\rho_p(S^{(1)}) < m^{1/p}$. With a Lipschitz - Linear prediction operator of order $N+1$, the convergence in $L^p(\mathbb{R}^d)$ is ensured provided $\rho_p(S^{(k)}) < m^{1/p}$, for some $k \leq N+1$.

The above theorem can easily by extended to a new inverse theorem in Besov spaces.

7. Stability in L^p and Besov Spaces

In applications, the multiscale data may be corrupted by some process. Since our model is nonlinear the inverse theorems does not ensure the stability. We develop here stability results for our new nonlinear formalism. To this end, we consider two data sets $(v^0, e^1, e^2, \cdots)$ and $(\tilde{v}^0, \tilde{e}^1, \tilde{e}^2, \cdots)$ corresponding to two reconstruction processes: $v^j = Sv^{j-1} + e^j$ and $\tilde{v}^j = S\tilde{v}^{j-1} + \tilde{e}^j$. In that context, we recall the definition of v as the limit of $v_j(x) = \sum_{k\in\mathbb{Z}^d} v_k^j \varphi_{j,k}(x)$, with $\varphi_{j,k}(x) = \varphi(M^j x - k)$ (and similarly for $\tilde{v}$). The stability of the multiscale representation in $L^p(\mathbb{R}^d)$, is stated by the following theorem:

Theorem 7.1 *Let S be a Lipschitz-Linear prediction operator of order $N+1$, and suppose that there exist an n and a $\rho < m^{1/p}$ such that:*

$$\|(S^{(k)})^n v - (S^{(k)})^n w\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}} \leq \rho^n \|v-w\|_{(\ell^p(\mathbb{Z}^d))^{r_k^d}} \ \forall v, w \in (\ell^p(\mathbb{Z}^d))^{r_k^d},$$

for some $k \leq N+1$. Assume that v_j and $\tilde{v}_j$ converge to v and $\tilde{v}$ in $L^p(\mathbb{R}^d)$ repsectively. Then, we have:

$$\|v-\tilde{v}\|_{L^p(\mathbb{R}^d)} \leq C\left(\|v^0-\tilde{v}^0\|_{L^p(\mathbb{R}^d)} + \sum_{l=1}^{j} m^{-l/p}\|e^l-\tilde{e}^l\|_{\ell^p(\mathbb{Z}^d)}\right) \text{ for some } C>0 \quad (7.1)$$

Analogously, we get a stability theorem in Besov space $B^s_{p,q}(\mathbb{R}^d)$.

8. $(\mathcal{A}, I)$-Compatible Nonlinear Prediction Operators

Given families of multi-indices I and of vectors $\mathcal{A}$, we define: $\Delta^{\mathcal{A}_I} = \left\{\Delta^{i_1}_{a_1}\cdots\Delta^{i_p}_{a_p},\ a_k \in \mathcal{A}, i_k \in I\right\}$. In other words, $\Delta^{\mathcal{A}_I}$ is a difference operator computed with respect to the family of vectors $\mathcal{A}$ and orders

given by I. Then, we have the definition of $(\mathcal{A}, I)$-compatible nonlinear prediction operator:
Definition 8.1 *A nonlinear prediction operator S is called $(\mathcal{A}, I)$-compatible if there exists a local linear prediction operator S_l and if it satisfies*

$$Sv_{Mk+i} = S_l v_{Mk+i} + \Phi_i(\Delta^{\mathcal{A}_I} v_{k+p_1}, \cdots, \Delta^{\mathcal{A}_I} v_{k+p_q}) \quad \forall i \in \ coset(M)$$

where $\{p_1, \cdots, p_q\}$ is a fix set, Φ_i are bounded Lipschitz functions and if there exists an operator $S_l^{\mathcal{A}_I}$ satisfying: $\Delta^{\mathcal{A}_I} S_l v = S_l^{\mathcal{A}_I} \Delta^{\mathcal{A}_I} v$.

From its definition, the operator S admits an operator $S^{\mathcal{A}_I}$. We also remark that *Lipschitz - Linear* operators of order $N+1$ are $(\mathcal{A}, I)$-compatible with $I = \{i; |i| = N+1\}$ and $\mathcal{A} = \{e_1, \ldots, e_d\}$. Note that we can extend all the notions described in the previous sections for *Lipschitz-Linear* prediction operators to $(\mathcal{A}, I)$-compatible prediction operators. The interest of using the notion of $(\mathcal{A}, I)$-compatibility is to provide simpler proofs of convergence and stability, as illustrated in the next section. Indeedb, the $(\mathcal{A}, I)$-compatibility enables to significantly reduce the number of computed differences to prove convergence and stability.

9. Applications

We study the convergence of bidimensional PPH multi-scale representations with prediction operator. We modify the PPH scheme as follows $\hat{v}^j_{Mk+e_1} = \frac{v_k^{j-1} + v_{k+Me_1}^{j-1}}{2} - \frac{\omega}{8} H(\Delta^2_{Me_1} v_k^{j-1}, \Delta^2_{Me_1} v_{k-Me_1}^{j-1})$ and $\hat{v}^j_{Mk} = v_k^{j-1}$, for some $0 < \omega < 1$. This prediction operator is *Lipschitz-Linear*, we also notice that it is $(\mathcal{A}, I)$-compatible with $\mathcal{A} = \{e_1, Me_1\}$ and $I = \{(0,2), (2,0)\}$, where M is the quincunx matrix. Therefore, to prove the convergence we only study the joint spectral radius of $S^{\mathcal{A}_I}$. To this end, we first compute the differences of order 2 in the directions $\{e_1, Me_1\}$. We then compute an upper bound for $\rho_p(S^{\mathcal{A}_I})$. Doing so, we are able to prove that the associated multiscale representation is bounded in L^∞ as soon as $\omega < 1$, and convergent in L^p for any $p \geq 1$ and $\omega = 1$. Using some properties of the Lipschitz function H, we are also able to prove the stability as soon as $p \geq 1$ and $\omega < 1/2$.

References

[1] S. Amat, R. Donat, J. Liandrat, J.C. Trillo, Analysis of a New Non-Linear Subdivision Scheme: Applications to Image Processing, Found. Comput. Math., 6, 193-226 (2006).

[2] V. Chappelier and C. Guillemot, Oriented Wavelet for Image Compression and Denoising, IEEE Transactions on Image Processing, 15, 2892-2903, (2006).

[3] B. Mateï, Smoothness Characterization and Stability in Nonlinear multi-scale Framework: Theoretical Results Asymptotic Analysis, 46, 277-309, (2005).

[4] B. Mateï and S. Meignen, A New Formalism for Nonlinear and Non-Separable Multiscale Representation, submitted.