

Prefilter Design for Forensic Resampling Estimation

David Vázquez-Padín¹, Fernando Pérez-González^{1,2,3}

1. Signal Theory and Communications Department, University of Vigo, Vigo 36310
2. Electrical and Computer Engineering Department, University of New Mexico, Albuquerque
3. GRADIANT, Vigo, Spain.

Foz do Iguaçu, Brazil. November 30, 2011.

Outline

- 1 Introduction
- 2 Problem statement and model description
- 3 Prefilter Design subject to the objective function
- 4 Experimental verification
- 5 Conclusions

Introduction

Problem:

- The presence of forged images in the news, in magazines or flowing through the Internet has become something prevalent these days.

Introduction

Problem:

- The presence of forged images in the news, in magazines or flowing through the Internet has become something prevalent these days.

Spatial transformations $\Rightarrow$ resampling operation:



(a) Original image 1



(b) Original image 2



(c) Tampered image

Introduction

Spatial transformations $\Rightarrow$ resampling operation:



(a) Original image 1



(b) Original image 2



(c) Tampered image

Possible solution: Detection of resampling factor inconsistencies

- **Popescu & Farid (2005)**: measure the correlation between pixels performing an EM algorithm and expose the periodic patterns.
- **Kirchner (2008)**: use the variance of prediction residuals to describe periodic artifacts.

Introduction

Possible solution: Detection of resampling factor inconsistencies

- **Popescu & Farid (2005)**: measure the correlation between pixels performing an EM algorithm and expose the periodic patterns.
- **Kirchner (2008)**: use the variance of prediction residuals to describe periodic artifacts.
- **Gallagher (2005)**: apply a differentiator to the resampled signal to enhance the spectrum peaks of the covariance.
- **Mahdian & Saic (2008)**: applies a differentiator to the image, use a Radon transform and then process the unidimensional covariance.
- **Vázquez-Padín et al. (2010)**: apply a two-dimensional statistical test for studying the presence of cyclostationarity.

Introduction

Possible solution: Detection of resampling factor inconsistencies

- **Popescu & Farid (2005)**: measure the correlation between pixels performing an EM algorithm and expose the periodic patterns.
- **Kirchner (2008)**: use the variance of prediction residuals to describe periodic artifacts.
- **Gallagher (2005)**: apply a differentiator to the resampled signal to enhance the spectrum peaks of the covariance.
- **Mahdian & Saic (2008)**: applies a differentiator to the image, use a Radon transform and then process the unidimensional covariance.
- **Vázquez-Padín et al. (2010)**: apply a two-dimensional statistical test for studying the presence of cyclostationarity.

Important remark:

Derivative prefilters are used as a way of boosting resampling traces.

Introduction

A question arises:

Which is the optimum prefilter for resampling estimation?

Introduction

A question arises:

Which is the optimum prefilter for resampling estimation?

Previous Work:

- **Dalgaard et al. (2010)**: show that derivative prefilters are a nearly optimum solution for asymptotically large resampling factor.

Introduction

A question arises:

Which is the optimum prefilter for resampling estimation?

Previous Work:

- **Dalgaard et al. (2010):** show that derivative prefilters are a nearly optimum solution for asymptotically large resampling factor.

Our contributions:

- Analytical framework based on cyclostationarity theory, using a 1-D first-order AR process to model a natural image.
- Cost function that measures the degree of detectability.
- Design of optimum prefilters for a particular resampling factor.
- Experimental validation in real scenarios.

Problem statement

Resampling process in the 1-D space:

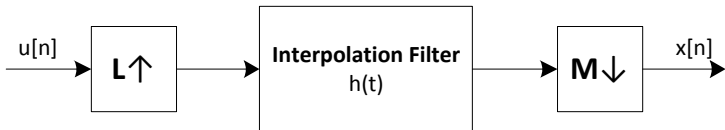


Figure: Block diagram representation of the resampling process.

Problem statement

Resampling process in the 1-D space:

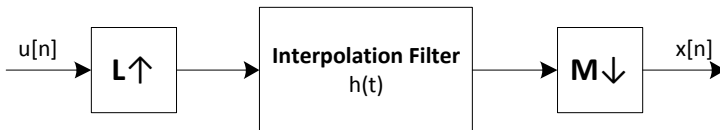


Figure: Block diagram representation of the resampling process.

Considerations:

- Resampling factor $N_s = \frac{L}{M}$ satisfies $1 < N_s < 2$.

Problem statement

Resampling process in the 1-D space:

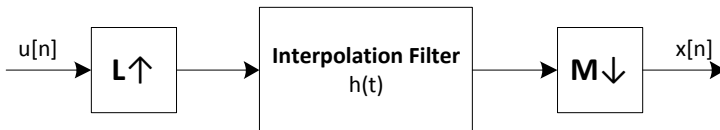


Figure: Block diagram representation of the resampling process.

Considerations:

- Resampling factor $N_s = \frac{L}{M}$ satisfies $1 < N_s < 2$.
- L and M are coprimes.

Problem statement

Resampling process in the 1-D space:

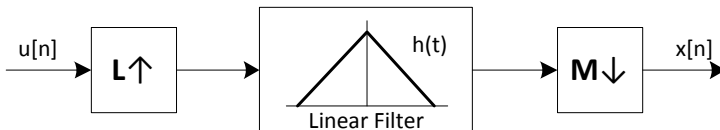


Figure: Block diagram representation of the resampling process.

Considerations:

- Resampling factor $N_s = \frac{L}{M}$ satisfies $1 < N_s < 2$.
- L and M are coprimes.
- Interpolation filter is linear.

Problem statement

Resampling process in the 1-D space:

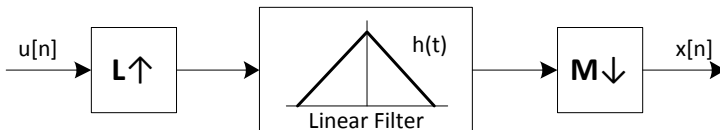


Figure: Block diagram representation of the resampling process.

Considerations:

- Resampling factor $N_s = \frac{L}{M}$ satisfies $1 < N_s < 2$.
- L and M are coprimes.
- Interpolation filter is linear.
- Input signal is wide-sense stationary.

Problem statement (II)

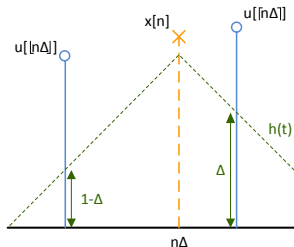


Figure: Illustrative example.

Resampling step: $\Delta = \frac{M}{L}$.

Interpolator:

$$h(t) = \begin{cases} 1 - |t|, & \text{if } |t| \leq 1 \\ 0, & \text{otherwise} \end{cases}.$$

Resampled value:

$$x[n] = u[\lfloor n\Delta \rfloor] (1 - \Delta) + u[\lceil n\Delta \rceil] \Delta.$$

Problem statement (II)

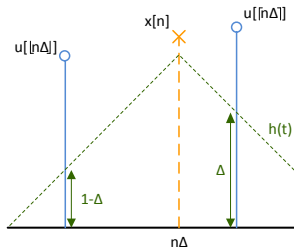


Figure: Illustrative example.

Resampling step: $\Delta = \frac{M}{L}$.

Interpolator:

$$h(t) = \begin{cases} 1 - |t|, & \text{if } |t| \leq 1 \\ 0, & \text{otherwise} \end{cases}.$$

Resampled value:

$$x[n] = u[\lfloor n\Delta \rfloor] (1 - \Delta) + u[\lceil n\Delta \rceil] \Delta.$$

General expression of the resampled signal ($v[n] = \text{mod}(n\Delta, 1)$):

$$x[n] = u[\lfloor n\Delta \rfloor] (1 - v[n]) + u[\lceil n\Delta \rceil] v[n].$$

Problem statement (III)

Covariance (correlation) of the zero-mean resampled signal $x[n]$:

$$c_{xx}[n; \tau] = E\{x[n]x[n + \tau]\}$$

Problem statement (III)

Covariance (correlation) of the zero-mean resampled signal $x[n]$:

$$\begin{aligned} c_{xx}[n; \tau] = & E \{ u[\lfloor n\Delta \rfloor] u[\lfloor (n + \tau)\Delta \rfloor] \} (1 - v[n]) (1 - v[n + \tau]) \\ & + E \{ u[\lfloor n\Delta \rfloor] u[\lceil (n + \tau)\Delta \rceil] \} (1 - v[n]) v[n + \tau] \\ & + E \{ u[\lceil n\Delta \rceil] u[\lfloor (n + \tau)\Delta \rfloor] \} v[n] (1 - v[n + \tau]) \\ & + E \{ u[\lceil n\Delta \rceil] u[\lceil (n + \tau)\Delta \rceil] \} v[n] v[n + \tau]. \end{aligned}$$

The resampled signal is wide-sense almost cyclostationary if:

$$c_{xx}[n; \tau] = c_{xx} \left[n + k \frac{L}{M}; \tau \right].$$

Problem statement (III)

Covariance (correlation) of the zero-mean resampled signal $x[n]$:

$$\begin{aligned} c_{xx}[n; \tau] = & E \{ u[\lfloor n\Delta \rfloor] u[\lfloor (n + \tau)\Delta \rfloor] \} (1 - v[n]) (1 - v[n + \tau]) \\ & + E \{ u[\lfloor n\Delta \rfloor] u[\lceil (n + \tau)\Delta \rceil] \} (1 - v[n]) v[n + \tau] \\ & + E \{ u[\lceil n\Delta \rceil] u[\lfloor (n + \tau)\Delta \rfloor] \} v[n] (1 - v[n + \tau]) \\ & + E \{ u[\lceil n\Delta \rceil] u[\lceil (n + \tau)\Delta \rceil] \} v[n] v[n + \tau]. \end{aligned}$$

The resampled signal is wide-sense almost cyclostationary if:

$$c_{xx}[n; \tau] = c_{xx} \left[n + k \frac{L}{M}; \tau \right].$$

Be aware that:

Although input signal $u[n]$ is **wide-sense stationary**, output signal $x[n]$ is **wide-sense almost cyclostationary** with period $N_s = \frac{L}{M}$.

Problem statement (III)

Covariance (correlation) of the zero-mean resampled signal $x[n]$:

$$\begin{aligned} c_{xx}[n; \tau] = & E \{ u[\lfloor n\Delta \rfloor] u[\lfloor (n + \tau)\Delta \rfloor] \} (1 - v[n]) (1 - v[n + \tau]) \\ & + E \{ u[\lfloor n\Delta \rfloor] u[\lceil (n + \tau)\Delta \rceil] \} (1 - v[n]) v[n + \tau] \\ & + E \{ u[\lceil n\Delta \rceil] u[\lfloor (n + \tau)\Delta \rfloor] \} v[n] (1 - v[n + \tau]) \\ & + E \{ u[\lceil n\Delta \rceil] u[\lceil (n + \tau)\Delta \rceil] \} v[n] v[n + \tau]. \end{aligned}$$

The resampled signal is wide-sense almost cyclostationary if:

$$c_{xx}[n; \tau] = c_{xx} \left[n + k \frac{L}{M}; \tau \right].$$

Important:

Due to this periodicity, we can provide an estimation of the resampling factor $\hat{N}_s$, since in the Fourier domain $C_{xx}(\omega; \tau) = \text{DFT}\{c_{xx}[n; \tau]\}$ there will be peaks at

$$\omega_1 = 2\pi \frac{M}{L} \quad \text{and} \quad \omega_2 = 2\pi \frac{L-M}{L} \quad (\text{due to symmetry}).$$

Model description

First-order autoregressive model:

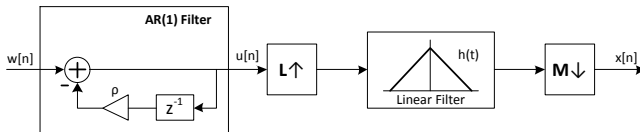


Figure: Block diagram representation of the whole model.

Example of modeling of images

First-order autoregressive model:



Figure: Real image.

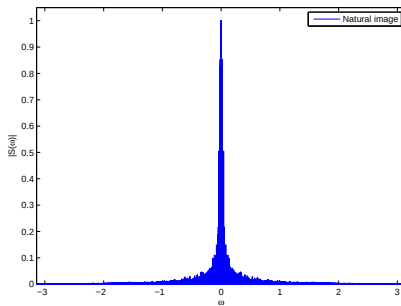


Figure: Spectral density.

Example of modeling of images

First-order autoregressive model:



Figure: Real image.

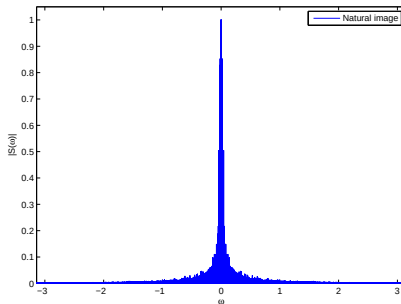


Figure: Spectral density.

Extracted AR(1) model parameters:

- $\rho = 0.9596$
- $\sigma_w^2 = 261$

Example of modeling of images

First-order autoregressive model:



Figure: Real image.

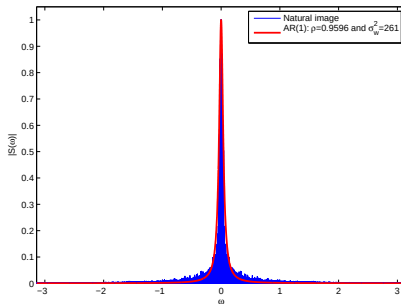


Figure: Spectral density comparison.

Extracted AR(1) model parameters:

- $\rho = 0.9596$
- $\sigma_w^2 = 261$

Model description

First-order autoregressive model:

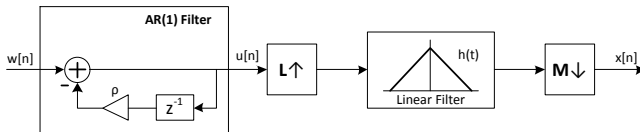


Figure: Block diagram representation of the whole model.

Model description

First-order autoregressive model:

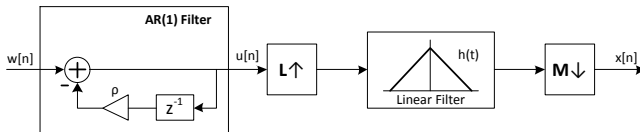


Figure: Block diagram representation of the whole model.

- **Input signal:** $u[n] = w[n] + \rho u[n - 1]$, with $w[n] \sim \mathcal{N}(0, 1)$.

Model description

First-order autoregressive model:

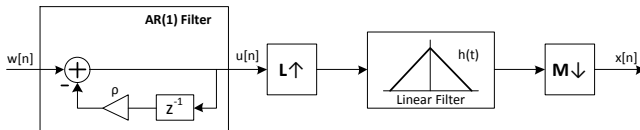


Figure: Block diagram representation of the whole model.

- **Input signal:** $u[n] = w[n] + \rho u[n-1]$, with $w[n] \sim \mathcal{N}(0, 1)$.
- **Correlation:** $c_{uu}[n; \tau] = E\{u[n]u[n+\tau]\} = \frac{\rho^{|\tau|}}{1-\rho^2}$.

Model description

First-order autoregressive model:

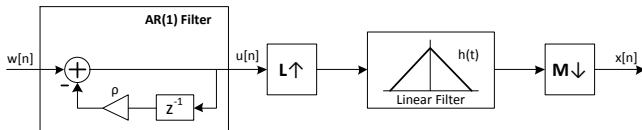


Figure: Block diagram representation of the whole model.

• Output correlation:

$$\begin{aligned}
 c_{xx}[n; \tau] = & \frac{1}{1 - \rho^2} \left[\rho^{| \lfloor n\Delta \rfloor - \lfloor (n+\tau)\Delta \rfloor |} (1 - v[n]) (1 - v[n + \tau]) \right. \\
 & + \rho^{| \lfloor n\Delta \rfloor - \lceil (n+\tau)\Delta \rceil |} (1 - v[n\Delta]) v[n + \tau] \\
 & + \rho^{| \lceil n\Delta \rceil - \lfloor (n+\tau)\Delta \rfloor |} v[n] (1 - v[n + \tau]) \\
 & \left. + \rho^{| \lceil n\Delta \rceil - \lceil (n+\tau)\Delta \rceil |} v[n] v[n + \tau] \right].
 \end{aligned}$$

Model description

First-order autoregressive model:

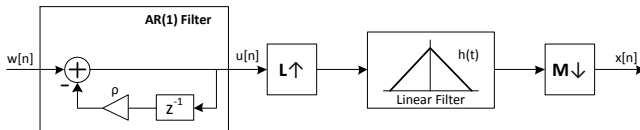


Figure: Block diagram representation of the whole model.

• Output correlation:

$$\begin{aligned}
 c_{xx}[n; \tau] = & \frac{1}{1 - \rho^2} \left[\rho^{|\lfloor n\Delta \rfloor - \lfloor (n+\tau)\Delta \rfloor|} (1 - v[n]) (1 - v[n + \tau]) \right. \\
 & + \rho^{|\lfloor n\Delta \rfloor - \lceil (n+\tau)\Delta \rceil|} (1 - v[n\Delta]) v[n + \tau] \\
 & + \rho^{|\lceil n\Delta \rceil - \lfloor (n+\tau)\Delta \rfloor|} v[n] (1 - v[n + \tau]) \\
 & \left. + \rho^{|\lceil n\Delta \rceil - \lceil (n+\tau)\Delta \rceil|} v[n] v[n + \tau] \right] = c_{xx} \left[n + k \frac{L}{M}; \tau \right].
 \end{aligned}$$

Example for $N_s = 11/10$

Parameters: $\rho = -0.95$, $L = 11$, $M = 10$ and $\tau = 0$.

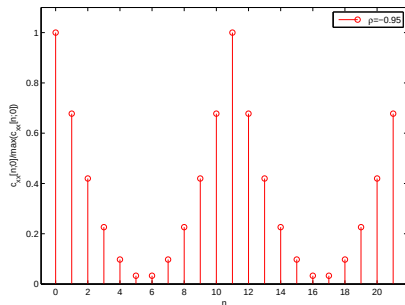


Figure: Correlation $c_{xx}[n; \tau]$.

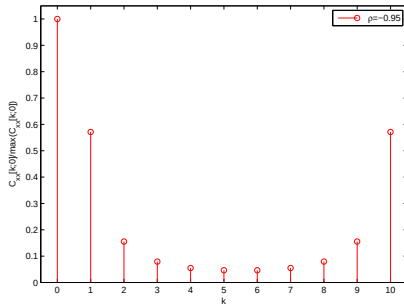


Figure: DTFS $C_{xx}[2\pi \frac{k}{L}; \tau]$.

Example for $N_s = 11/10$

Parameters: $\rho \approx 0$, $L = 11$, $M = 10$ and $\tau = 0$.

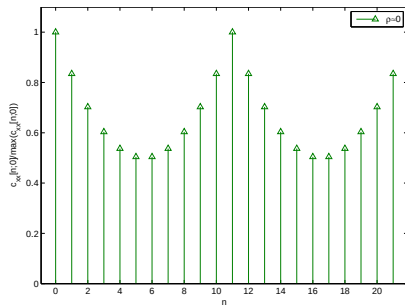


Figure: Correlation $c_{xx}[n; \tau]$.

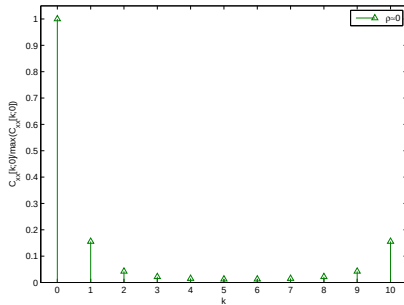


Figure: DTFS $C_{xx}\left[2\pi\frac{k}{L}; \tau\right]$.

Example for $N_s = 11/10$

Parameters: $\rho = 0.95$, $L = 11$, $M = 10$ and $\tau = 0$.

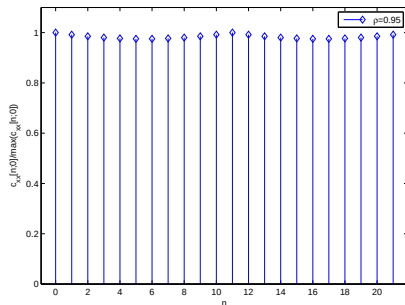


Figure: Correlation $c_{xx}[n; \tau]$.

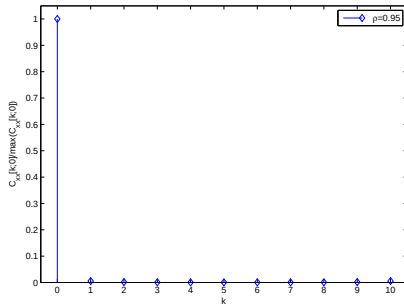


Figure: DTFS $C_{xx}\left[2\pi\frac{k}{L}; \tau\right]$.

Example for $N_s = 11/10$

Parameters: $\rho = 0.95$, $L = 11$, $M = 10$ and $\tau = 0$.

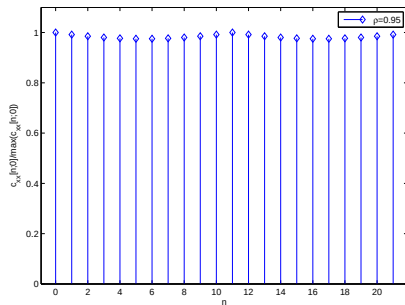


Figure: Correlation $c_{xx}[n; \tau]$.

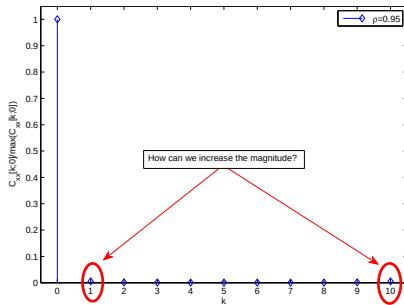


Figure: DTFS $C_{xx}\left[2\pi\frac{k}{L}; \tau\right]$.

Prefilter Design

P-order prefilter with coefficients p_l with $l \in \{0, \dots, P\}$:

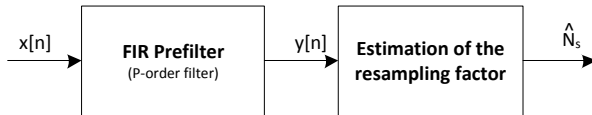


Figure: Location of the prefilter.

Prefilter Design

P-order prefilter with coefficients p_l with $l \in \{0, \dots, P\}$:

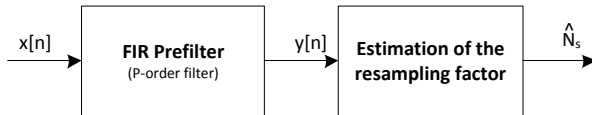


Figure: Location of the prefilter.

Output correlation:

$$c_{yy}[n; \tau] = \sum_{l=0}^P \sum_{m=0}^P p_l p_m c_{xx}[n - l; \tau + l - m].$$

Prefilter Design

P-order prefilter with coefficients p_l with $l \in \{0, \dots, P\}$:

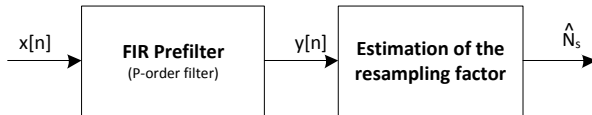


Figure: Location of the prefilter.

Objective function:

$$\Theta(L, M, \rho, p_l) \doteq \frac{\frac{1}{2}(|C_{yy}[2\pi \frac{M}{L}; 0]|^2 + |C_{yy}[2\pi \frac{L-M}{L}; 0]|^2)}{\frac{1}{L-2} \sum_{\substack{k=0 \\ k \neq M, L-M}}^{L-1} |C_{yy}[2\pi \frac{k}{L}; 0]|^2}$$

Cyclic frequencies:

$$\alpha_1 = 2\pi \frac{M}{L}, \quad \alpha_2 = 2\pi \frac{L-M}{L}$$

Evaluation of the cost function without prefilter

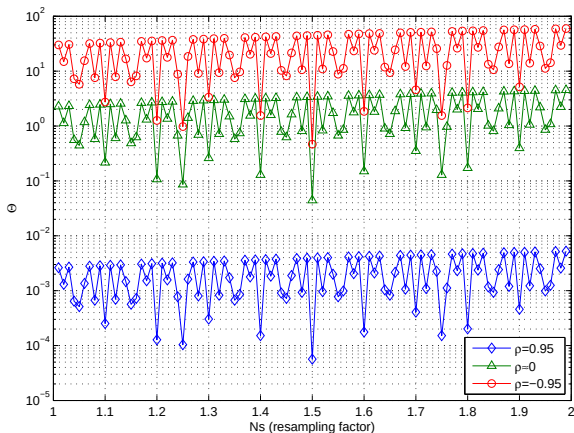


Figure: Objective function for resampling factors in $1 < N_s < 2$ and for different values of ρ .

Evaluation of the cost function with prefilter (I)

Parameters: $\rho = 0.95$ and $N_s = 11/10$.

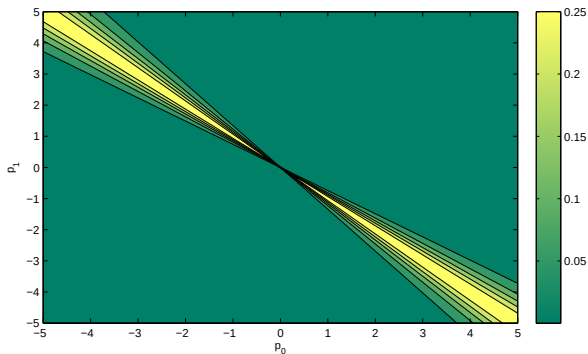


Figure: Objective function Θ considering a first-order prefilter, varying the coefficients p_0 and p_1 in the range $[-5, 5]$.

Evaluation of the cost function with prefilter (II)

Parameters: $\rho = 0.95$ and $N_s = 11/10$.

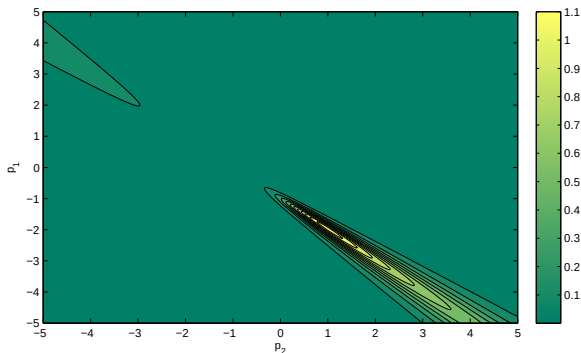


Figure: Objective function Θ considering a second-order prefilter, fixing $p_0 = 1$ and varying the coefficients p_1 and p_2 in the range $[-5, 5]$.

Evaluation of the cost function with prefilter (III)

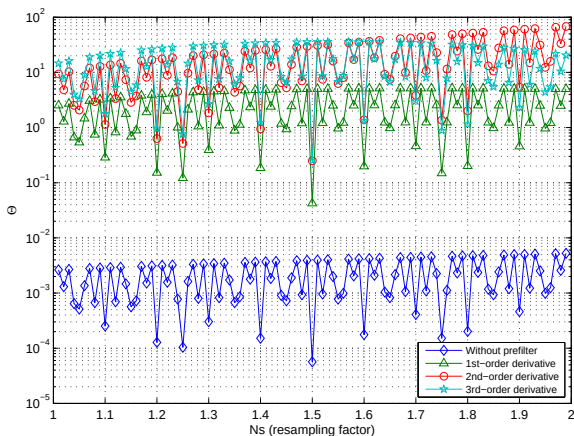


Figure: Objective function for resampling factors in $1 < N_s < 2$ and for $\rho = 0.95$.

Evaluation of the cost function with prefilter (III)

Table: Optimum prefilters of order 3 for some values of N_s .

Range of N_s	Coefficients of the prefilter			
	p_0	p_1	p_2	p_3
1.05 - 1.1	1	-2.4	2.4	-1
1.3 - 1.35	1	-2.75	2.75	-1
1.4 - 1.45	1	-2.8	2.8	-1
1.6 - 1.65	1	-5	7.5	-3.5
1.85 - 1.95	1	-2	1.1429	-0.1429

Experimental verification

Estimation process:

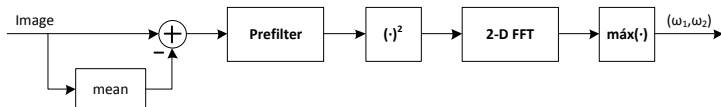


Figure: Block diagram representation of the estimation process.

Simulation details:

- Image database: 150 non-compressed RGB-images.
- Process only G component.
- Block-size of analysis: 1024×1024 pixels.
- Exclude DC component.
- If $\left| \max_{i \in \{1,2\}} |\omega_i| - \alpha \right| \leq \frac{2\pi}{1024}$, then $\hat{N}_s = \frac{2\pi}{2\pi - \max_{i \in \{1,2\}} |\omega_i|}$.

Experimental verification (II)

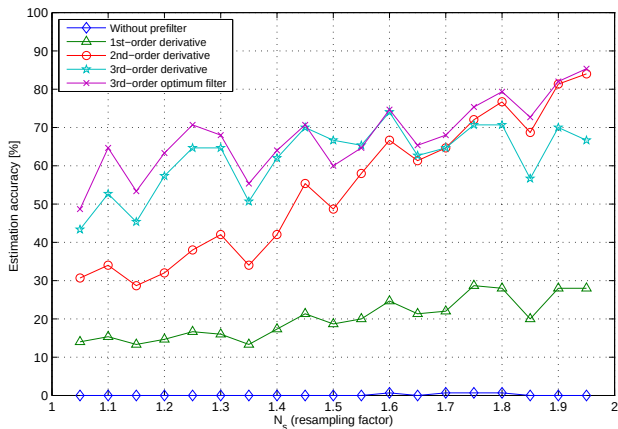


Figure: Estimation accuracy of the resampling factor.

Experimental verification (III)

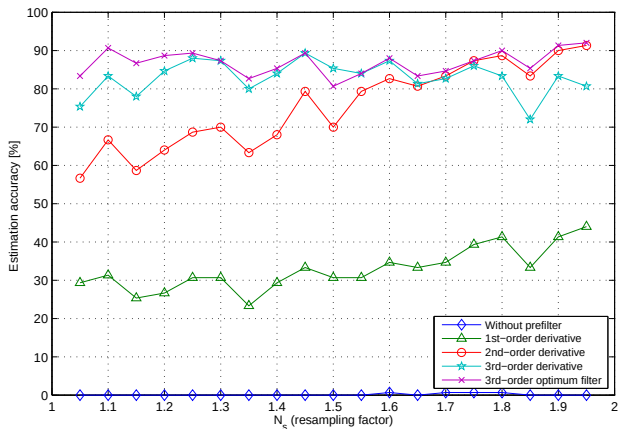


Figure: Estimation accuracy of the resampling factor, excluding near-zero frequencies.

Conclusions

Conclusions:

- Design of prefilters to improve the estimation accuracy has been analytically investigated.
- Experimental results validate the proposed model and the cost function.

Further work:

- Improvement of the cost function for the design of the prefilter by considering the rounding operation.
- Analysis of the windowing and his effect on the design of the optimum prefilter.
- Define a way to properly estimate the resampling factor when $N_s < 1$.
- Spread the study of the optimum prefilter for different interpolation filters.

Thanks for your attention.

Questions?

David Vázquez-Padín
dvazquez@gts.uvigo.es