Vector Lifting Schemes for Stereo Images Coding

M. Kaaniche[1], A. Benazza[2], B. Pesquet-Popescu[1] and J.-C. Pesquet[3]

[1] Telecom-ParisTech,
Département Signal et Images

[2] Ecole Supérieure des Communications de Tunis
Unité de Recherche en Imagerie Satellitaire et ses Applications (URISA)

[3] Université de Paris-Est, Marne-la-Vallée,
Laboratoire d'Informatique, Equipe Signal et Communications

October 23, 2008

Outline

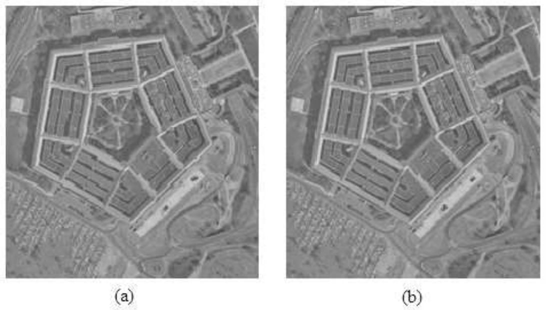
1. Context of study
2. Basic approach for joint coding of stereo image
3. Novel joint SI coding
4. Performances evaluation
5. Conclusions and perspectives

Part I

Context of study

Type of data

Stereo Image (SI): Two images, captured from two viewpoints, corresponding to the same scene



The SI pair “pentagon”: (a) left image, (b) right image

Interest of SI:

3D shape reconstruction in remote sensing, medical imaging

3D reconstruction of the SI pair “pentagon”

Problem:

Constitution of enormous amounts of data.

Our objectives: design a Stero Image (SI) coding scheme with

- Lossless reconstruction
  $\Longrightarrow$ exact decoding of SI (required for remote sensing imaging applications)
- Progressive reconstruction
  $\Longrightarrow$ gradual decoding that generates two compact multiresolution representations of SI (suitable for telebrowsing applications)

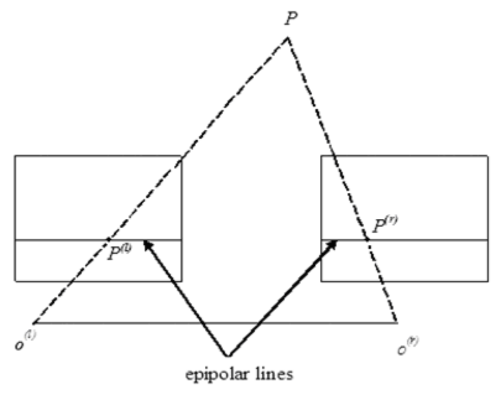
Part II

Basic approach for joint coding of stereo image

Croos-view redundancies

Binocular imaging system:

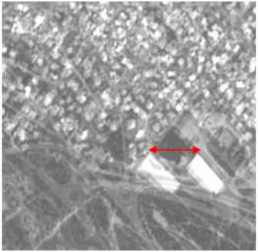
Assumption: rectified images.

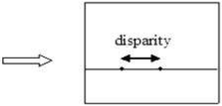


Binocular imaging system

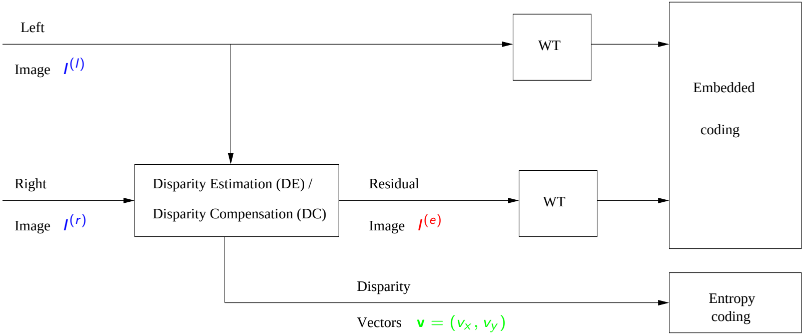
⇓

If the two images are superimposed





Generic decomposition scheme for SI progressive coding



$$\hat{I}^{(r)}(m,n) = I^{(l)}(m+v_x, n+v_y). \tag{1}$$

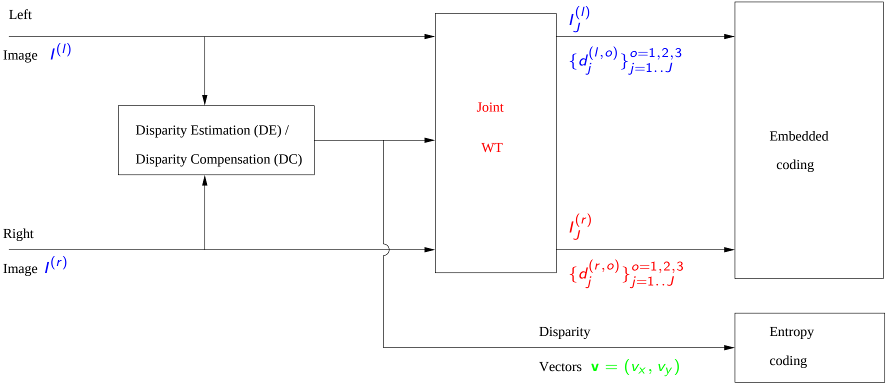
$$I^{(e)}(m,n) = I^{(r)}(m,n) - \hat{I}^{(r)}(m,n) \tag{2}$$

Part III

Novel joint SI coding

Motivation:

A new approach based on the Vector Lifting Scheme (VLS).



Novel joint SI coding

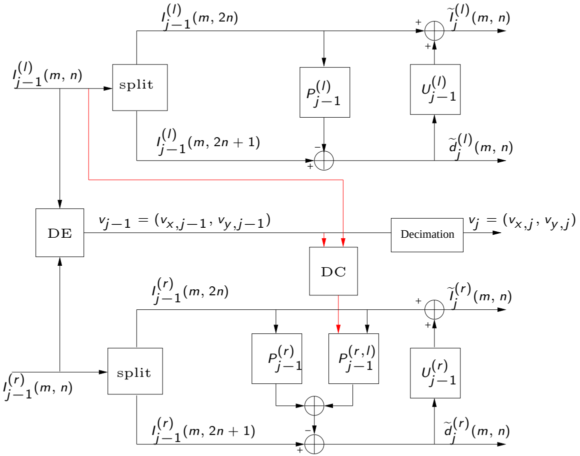
Advantages:

- No generation of a residual image, but two multiresolution representations of $I^{(l)}$ and $I^{(r)}$.
- Separable decompositions $\Longrightarrow$ Simplicity of their implementation

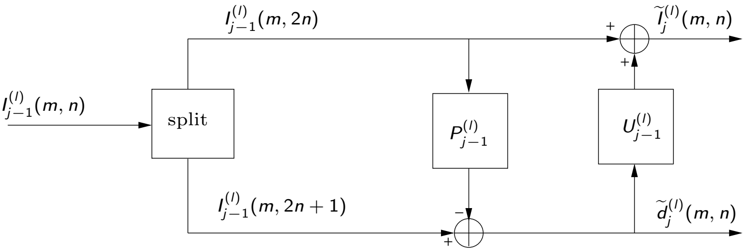
Two decomposition examples:

- VLS-I
- VLS-II

Principle of VLS-I decomposition



Equations of VLS-I

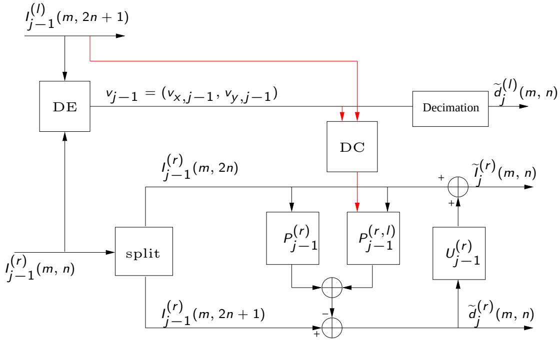


For the reference image $I^{(l)}$:

$$\tilde{d}_j^{(l)}(m,n) = I_{j-1}^{(l)}(m,2n+1) - \lfloor \frac{1}{2}(I_{j-1}^{(l)}(m,2n) + I_{j-1}^{(l)}(m,2n+2)) \rceil \quad (3)$$

$$\tilde{I}_j^{(l)}(m,n) = I_{j-1}^{(l)}(m,2n) + \lfloor \frac{1}{4}(\tilde{d}_j^{(l)}(m,n-1) + \tilde{d}_j^{(l)}(m,n)) \rceil \quad (4)$$

Equations of VLS-I



For the right image $I^{(r)}$:

$$\tilde{d}^{(r)}_j(m,n) = I^{(r)}_{j-1}(m,2n+1) - \lfloor p^{(r)}_{j-1,1} I^{(r)}_{j-1}(m,2n) + p^{(r)}_{j-1,2} I^{(r)}_{j-1}(m,2n+2)$$
$$+ p^{(r,l)}_{j-1,3} I^{(l)}_{j-1}(m + v_{x,j-1}(m,2n+1), 2n+1+v_{y,j-1}(m,2n+1))\rceil \quad (5)$$

$$\tilde{I}^{(r)}_j(m,n) = I^{(r)}_{j-1}(m,2n) + \lfloor \frac{1}{4}(\tilde{d}^{(r)}_j(m,n-1) + \tilde{d}^{(r)}_j(m,n))\rceil \quad (6)$$

An Improved version: VLS-II

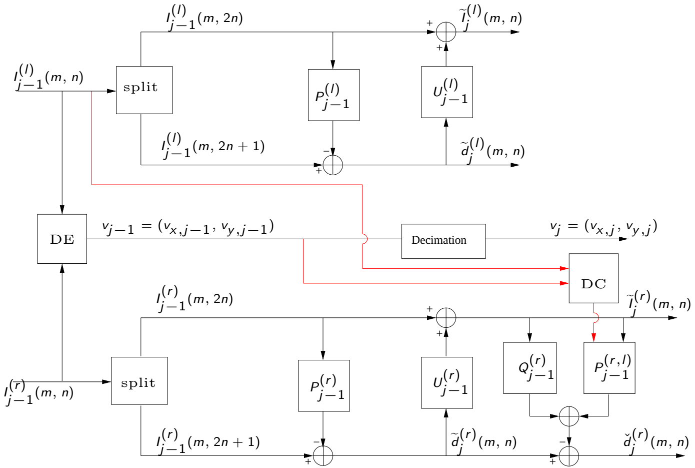
Motivation:

VLS-I: P-U structure $\Rightarrow$ approximation coefficients $\tilde{I}_j^{(l)}(m,n)$ inserted into $\tilde{d}_j^{(r)}(m,n)$ and then into $\tilde{I}_j^{(r)}(m,n)$ $\Rightarrow$ an update leakage effect.

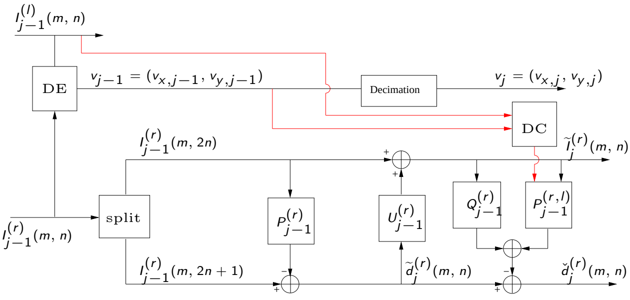
Proposed solution VLS-II: Another lifting with a P_1-U-P_2 structure.

- P_1 step: compute an intermediate detail signal by exploiting only the intra-image redundancies.
- U step: compute the approximation signal based on this intermediate detail signal.
- P_2 step: compute the final detail signal by exploiting the intra and inter-images redundancies.

Principle of VLS-II decomposition



Equations of VLS-II



$$\tilde{d}_j^{(r)}(m,n) = I_{j-1}^{(r)}(m,2n+1) - \lfloor \frac{1}{2}(I_{j-1}^{(r)}(m,2n) + I_{j-1}^{(r)}(m,2n+2))\rceil, \qquad (7)$$

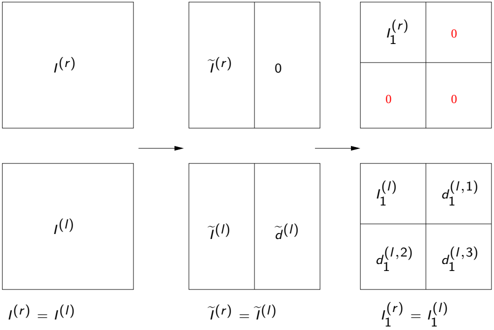
$$\tilde{I}_j^{(r)}(m,n) = I_{j-1}^{(r)}(m,2n) + \lfloor \frac{1}{4}(\tilde{d}_j^{(r)}(m,n-1) + \tilde{d}_j^{(r)}(m,n))\rceil, \qquad (8)$$

$$\check{d}_j^{(r)}(m,n) = \tilde{d}_j^{(r)}(m,n) - \lfloor q_{j-1}(\tilde{I}_j^{(r)}(m,n) + \tilde{I}_j^{(r)}(m,n+1)) + \sum_{k=-3}^{3} p_{j-1,k}^{(r,l)} s_{j-1}^{(l)}(m + v_{x,j-1}(m,2n+1), 2n+1+v_{y,j-1}(m,2n+1)-k)\rceil, \qquad (9)$$

Version 2: VLS-II

Advantage of VLS-II:

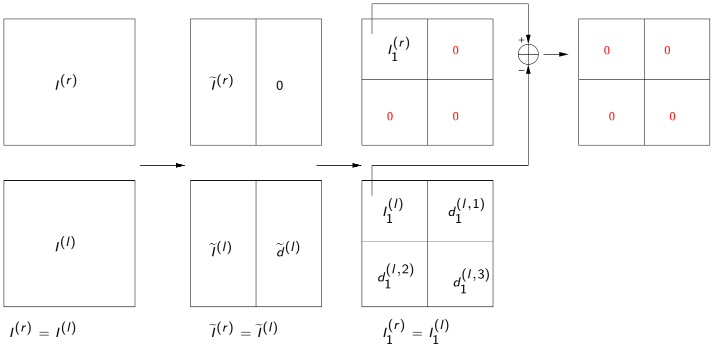
If $I^{(r)}=I^{(l)}$ and $\{0,1,-1\} \subseteq \mathcal{P}_j^{(r,l)} \Rightarrow \widetilde{I}^{(r)} = \widetilde{I}^{(l)}$ and $\check{d}^{(r)} = 0 \Rightarrow I_1^{(r)} = I_1^{(l)}$ and $d_1^{(r,o)} = 0, \quad o \in \{1,2,3\}$



Advantage of VLS-II:

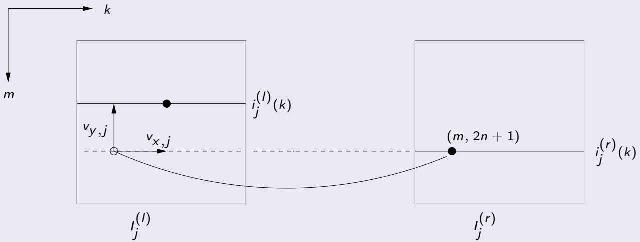
Finally, at the last resolution level $j = J$, instead of coding $I_J^{(r)}$, we code the residual subimage:

$$e_J^{(r)}(m,n) = I_J^{(r)}(m,n) - I_J^{(l)}(m + v_{x,J}(m,n), n + v_{y,J}(m,n)) \qquad (10)$$



Theoretical analysis of VLS-II:

Let $(m, 2n+1)$ be a given pixel to be predicted,



$$\left\{ \begin{array}{ll} i_j^{(r)}(k) & = I_j^{(r)}(m,k), \\ i_j^{(l)}(k) & = I_j^{(l)}(m+v_{x,j}(m,2n+1), k+v_{y,j}(m,2n+1)) \end{array} \right. , \qquad (11)$$

Theoretical analysis of VLS-I and VLS-II:

Assumption

$$\left\{\begin{array}{ll} i_j^{(r)}(k) & = \sin(\theta_j)x_j(k) + \cos(\theta_j)y_j(k) \\ i_j^{(I)}(k) & = \cos(\theta_j)x_j(k) + \sin(\theta_j)y_j(k) \end{array}\right. , \quad (12)$$

where

- x_j, y_j: AR(1), independent.
- $E[x_j(k)] = E[y_j(k)] = 0.$
- $E[\{x_j(k)\}^2] = E[\{y_j(k)\}^2] = \sigma_j^2.$

Theoretical analysis of VLS-I and VLS-II:

Some properties

- $E[x_j(k)x_j(k-l)] = E[y_j(k)y_j(k-l)] = \sigma_j^2 \rho_j^{|l|}$.
- $E\left[i_j^{(r)}(n) i_j^{(r)}(n-k)\right] = E\left[i_j^{(l)}(n) i_j^{(l)}(n-k)\right] = \rho_j^{|k|}$
- $E\left[i_j^{(r)}(n) i_j^{(l)}(n-k)\right] = s_j \rho_j^{|k|}$
  where $s_j = \sin(2\theta_j)$.
- The factor θ_j controls the cross-redundancies between the samples $i_j^{(l)}(k)$ and $i_j^{(r)}(k)$.

Minimum prediction error variance

- independent scheme:

$$E[\{\tilde{d}_j^{(r)}(k)\}^2] = \frac{1}{2}(1-\rho_j)(3-\rho_j) \tag{13}$$

- VLS-I:

$$\varepsilon_{1,j}(\rho_j, \theta_j) = \sigma_j^2 \gamma_{1,j} \cos^2(2\theta_j)(\rho_j^2 - 1) \tag{14}$$

where

$$\gamma_{1,j} = 2\sin^2(2\theta_j)(\rho_j^2 - \rho_j^2 - 1)^{-1}$$

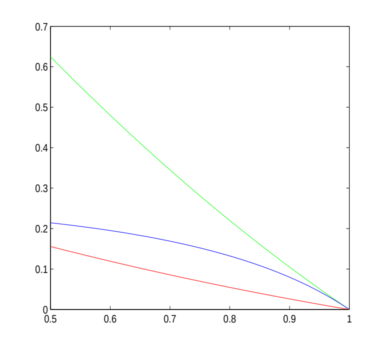
- VLS-II:

$$\varepsilon_{2,j}(\rho_j, \theta_j) = \frac{1}{2}\sigma_j^2 \gamma_{2,j} \cos^2(2\theta_j)(1-\rho_j)(3\rho_j^4 - 16\rho_j^3 + 4\rho_j^2 + 24\rho_j + 113) \tag{15}$$

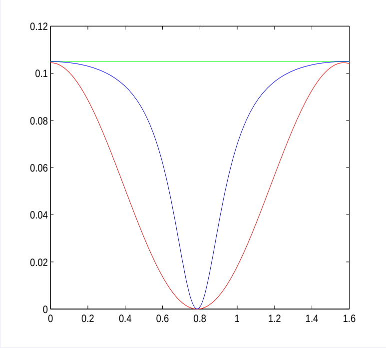
where $\gamma_{2,j} = (\rho_j^5 - 5\rho_j^4 - \rho_j^3 + 13\rho_j^2 + 18\rho_j + 38)^{-1}$.

Performances of independent scheme, VLS-I and VLS-II in terms of prediction effeciency

$E[\{\tilde{d}_j^{(}r)(n)\}^2]$
$\varepsilon_{1,j}(\rho_j,\theta_j)$
$\varepsilon_{2,j}(\rho_j,\theta_j)$



(a): Variations w.r.t ρ_j



(b): Variations w.r.t θ_j

Part IV

Performances evaluation

Performances evaluation

Experiments setup:

- Test images: natural and satellite (SPOT5) stereo images.
- Block size: 8×8.
- Search area S: [50,$\pm$2] for SPOT5 stereo images and [30,$\pm$4] for natural ones.
- Decomposition depth: $J = 2$.
- Performances evaluation in terms of bit rate, PSNR and SSIM (Structural SIMilarity)

Performances evaluation

Methods used for comparison:

- Independent scheme : Applying the 5/3 transform separately to original images $I^{(l)}$ and $I^{(r)}$.
- Scheme B : Applying the 5/3 transform to the reference and residual images $I^{(l)}$ and $I^{(e)}$.
- Scheme C (version of JPEG2000 scheme): Applying the 5/3 transform to $\tilde{I}^{(l)}$ and $I^{(e)}$, where

$$\begin{cases} I^{(e)}(m_x, m_y) & = I^{(r)}(m_x, m_y) - I^{(l)}(m_x + v_x, m_y + v_y) \\ \tilde{I}(m_x + v_x, m_y + v_y) & = \lfloor (I^{(r)}(m_x, m_y) + I^{(l)}(m_x + v_x, m_y + v_y))/2 \rfloor \\ \quad \text{if } (m_x + v_x, m_y + v_y) \in \mathcal{S} & \\ \tilde{I}(m_x, m_y) & = I^{(l)}(m_x, m_y) \qquad\qquad \text{if } (m_x, m_y) \notin \mathcal{S} \end{cases} \tag{16}$$

with $\mathcal{S}$ is the set of connected pixels in the left image.

For all methods, wavelets coefficients are encoded by applying the JPEG2000 codec.

PSNR curve

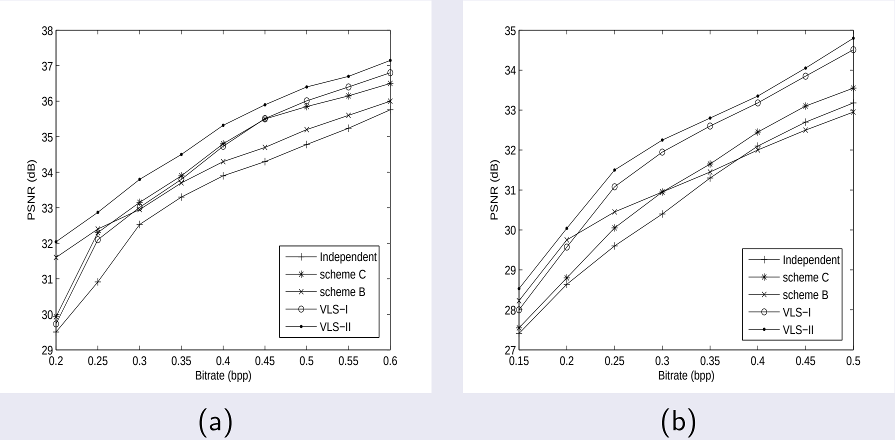


Figure: PSNR (in dB) versus the bitrate (bpp) after JPEG 2000 encoding for the SI pair “shrub” (a) and “spot5-6” (b).

Performances evaluation

Image visual quality (PSNR and SSIM)

(a): PSNR=30.03 dB, SSIM=0.80

(b): PSNR=31.48 dB, SSIM=0.83

Figure: Reconstructed target image $I^{(r)}$ of the "spot5-5" SI at 0.13 bpp: (a) scheme B; (b) VLS-II.

Final bitrate

Transform	scheme B	scheme C	VLS-I	VLS-II
spot5-1	3.63	3.58	**3.49**	**3.35**
spot5-2	3.85	3.78	**3.67**	**3.53**
spot5-3	4.27	4.24	**4.03**	**3.93**
spot5-4	4.22	4.21	**4.05**	**3.92**
spot5-5	3.91	3.89	**3.80**	**3.73**
spot5-6	3.89	3.81	**3.73**	**3.63**
fruit	4.05	3.97	**3.78**	**3.72**
shrub	3.73	3.69	**3.81**	**3.63**
birch	4.52	4.47	**4.44**	**4.37**
pentagon	5.37	5.2	**5.12**	**5.04**
Average	4.14	4.08	**3.99**	**3.88**

Conclusion and Perspectives

Presentation of two versions of a novel joint coding methods for stereo pairs.

$$\Downarrow$$

Improvement by designing more sophisticated prediction/update operators.