Optimal Dense Disparity Map Quantization and Residual Prediction for Lossless Stereo Image Coding

Presented by:

Amit Kumar K.C.

15th December 2010

Organization of presentation

- Objectives
- Introduction
- [Proposed scheme]
- Results
- Comparison with other schemes

Stereo images

Left image

Right image

- DISPARITY
  - Amount of shift needed such that a pixel on left image corresponds to that on right image

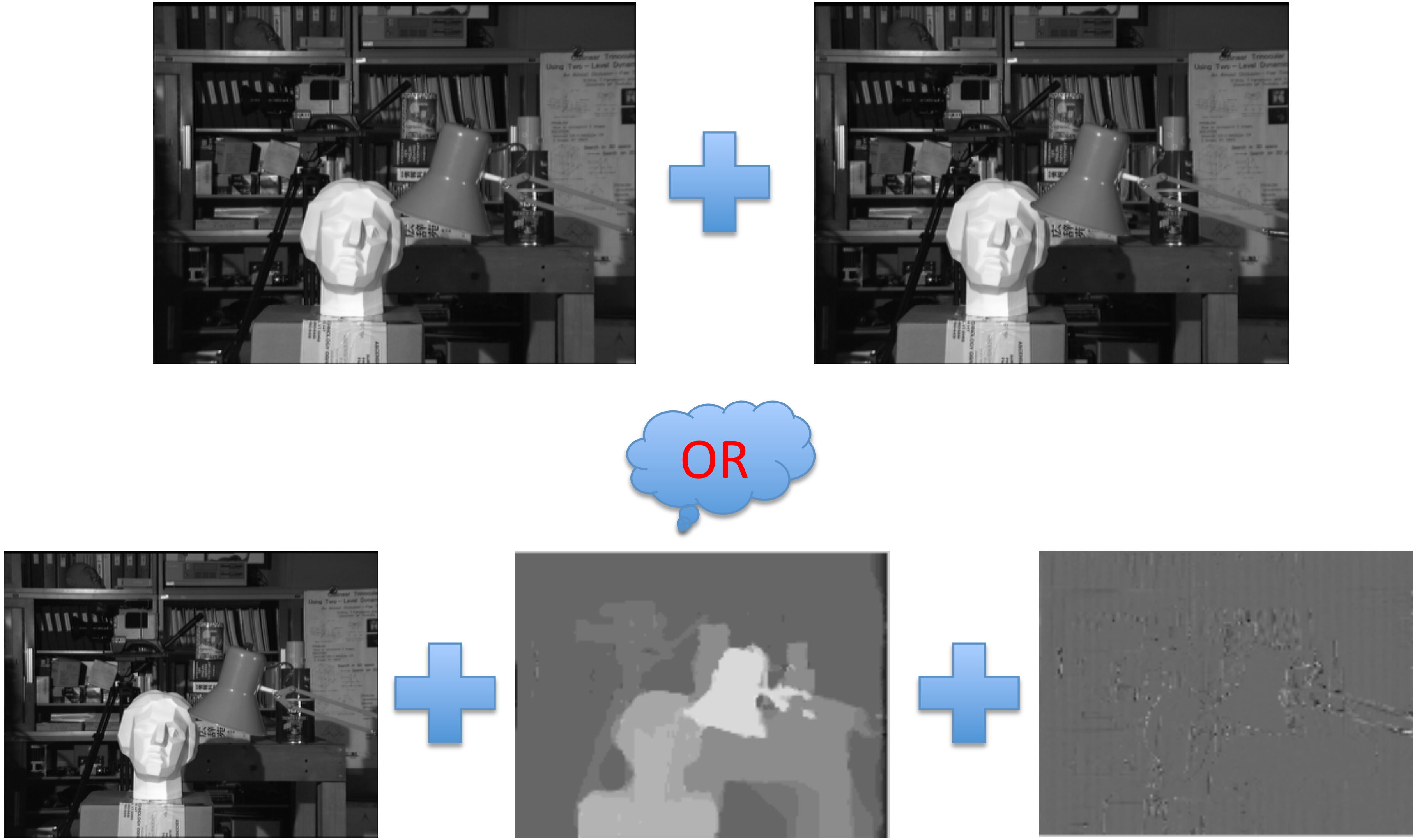
- RESIDUE
  - The difference between actual pixel intensity and disparity compensated pixel intensity

$$\underline{v}^*(\underline{x}) = \underset{\underline{v}(\underline{x})}{\arg\min}\, d\left(I^{(r)}(\underline{x}), I^{(l)}(\underline{x} + \underline{v}(\underline{x}))\right)$$

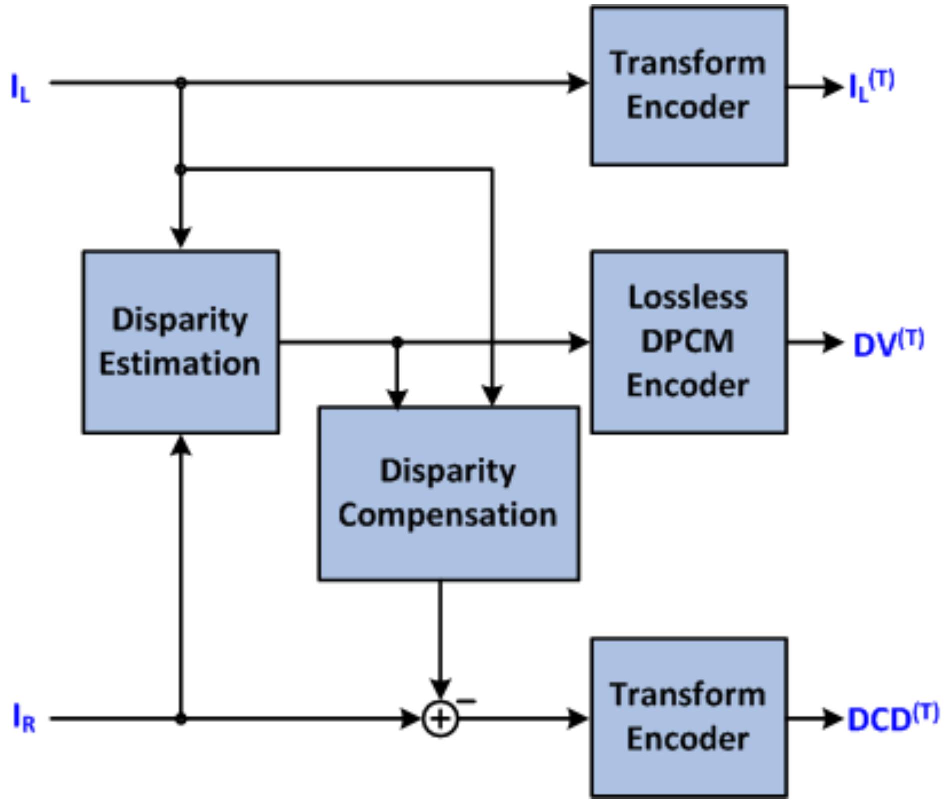
$$D(\underline{x}) = I^{(r)}(\underline{x}) - I^{(l)}(\underline{x} + \underline{v}^*(\underline{x}))$$

$I^{(r)}$ → right image
$I^{(l)}$ → left image
v* → disparity vector

d(a,b) → distance metric between a and b
D → Residual image

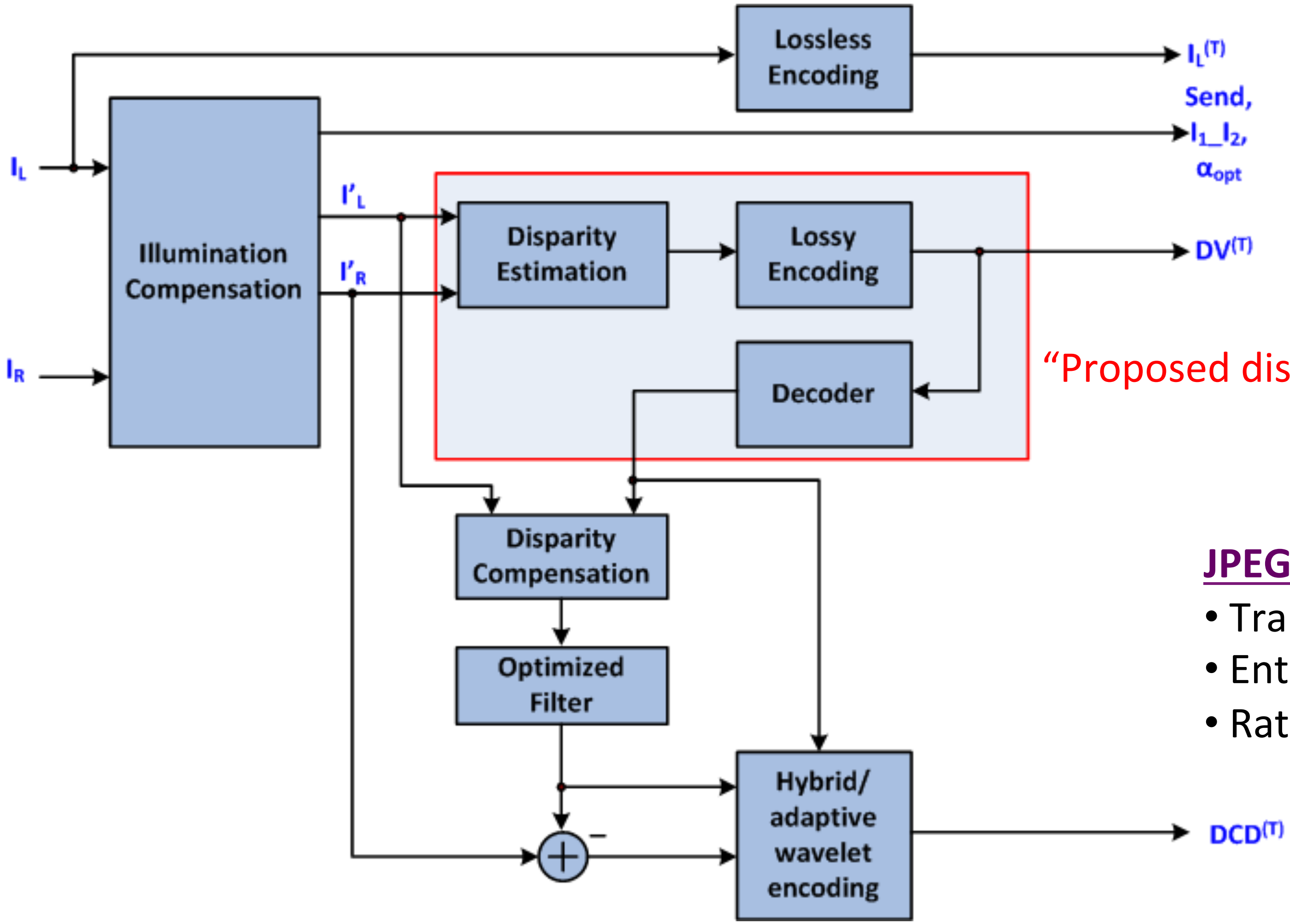

OR

Introduction



- **Block matching algorithm** for disparity estimation/compensation
- Transform coder: DCT, DWT
- DWT is preferred for digital cinema, HD images, medical images
- Normalized correlation → max, block illumination shift [Darazi09]

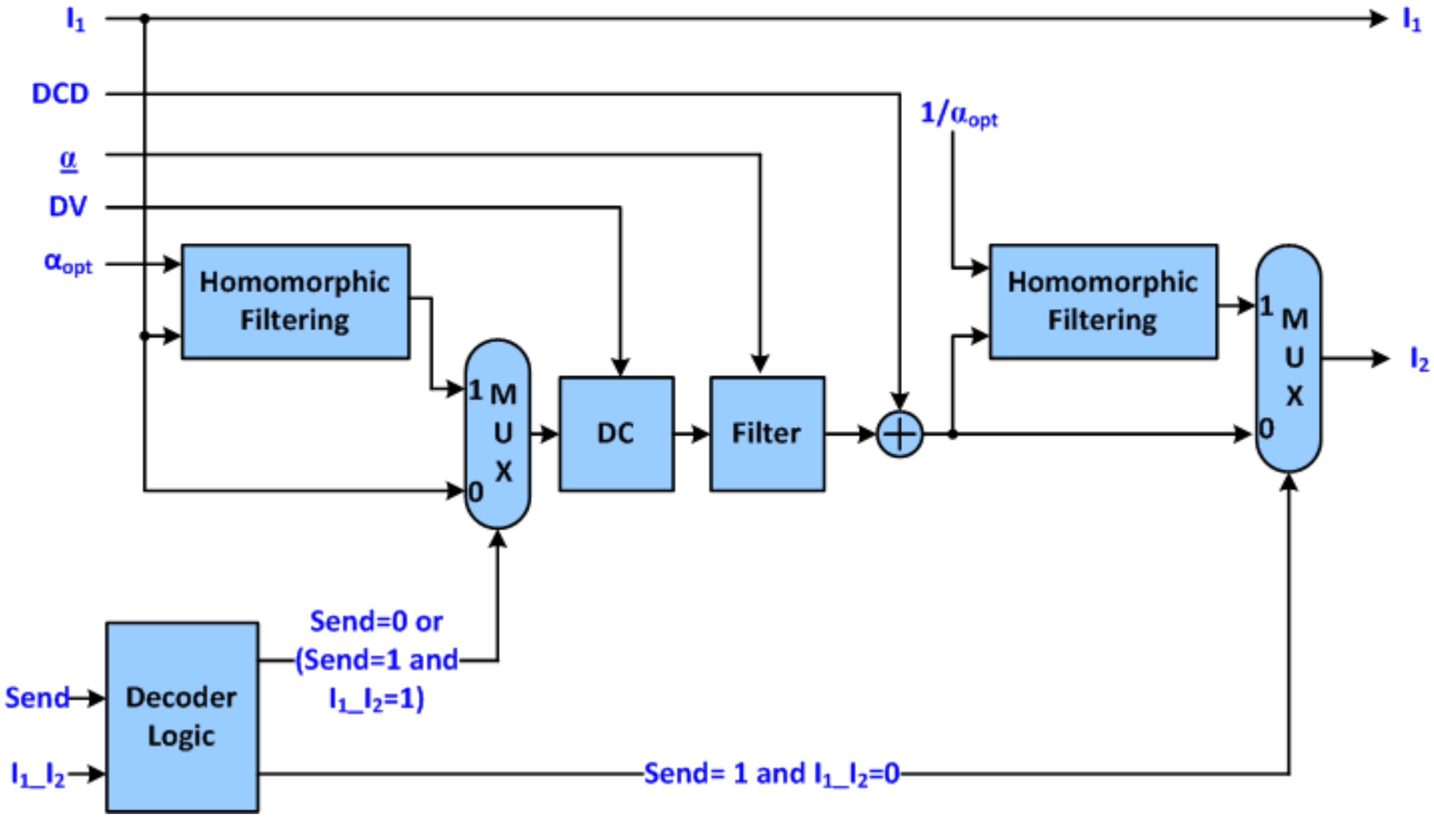
Proposed scheme-encoder



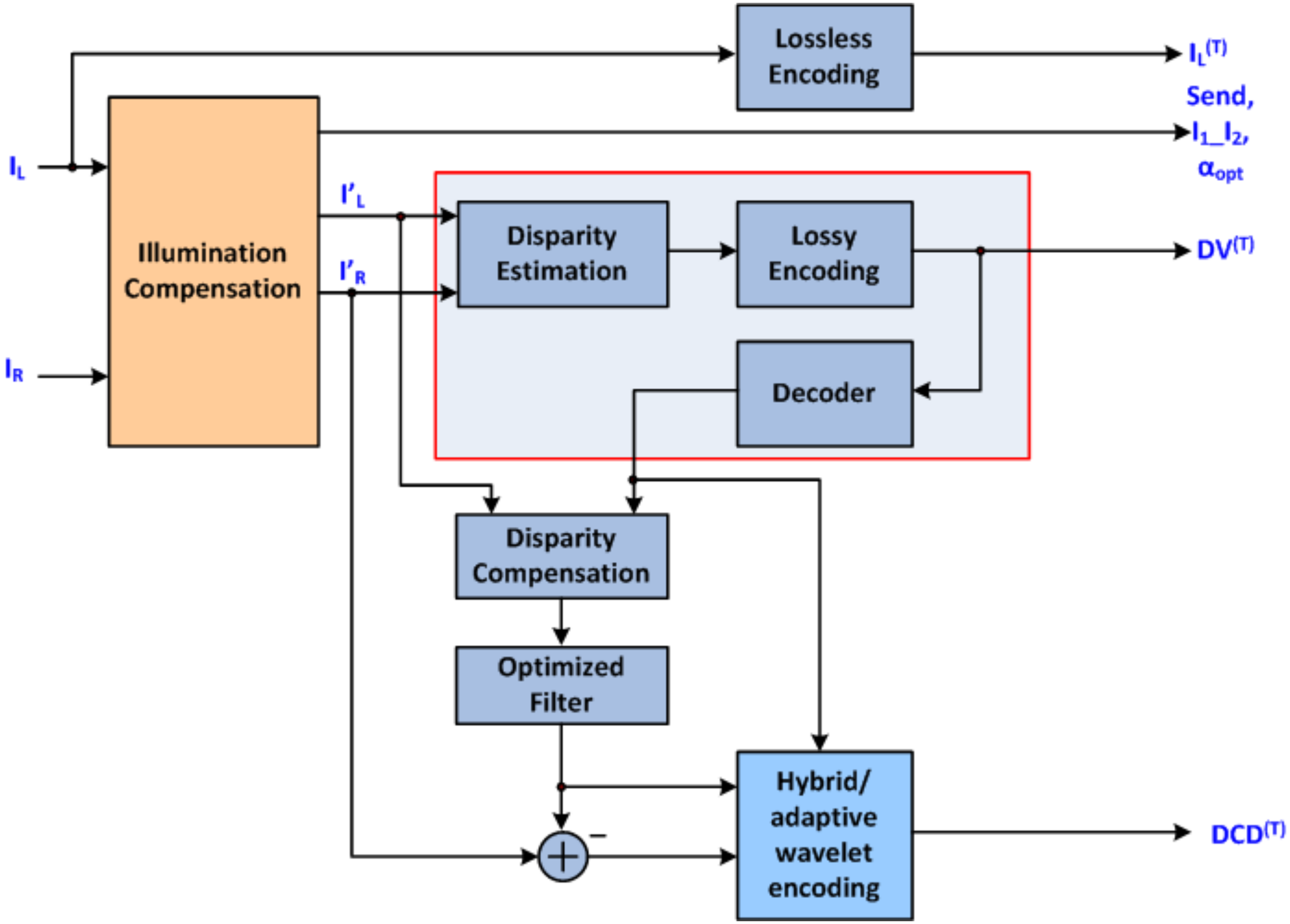
JPEG2000

- Transform coder
- Entropy coder
- Rate controller

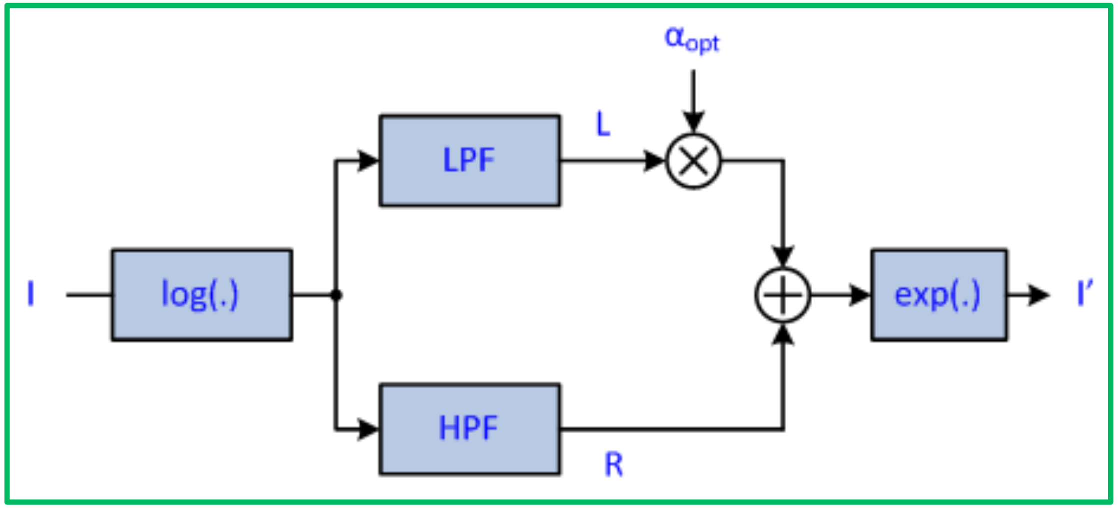
Proposed scheme-decoder



Illumination compensation



Illumination compensation



- Global illumination compensation
- **Homomorphic filter**
- Only one coefficient and two flag bits per image pair

$$I[m,n] = L[m,n] \times R[m,n]$$

Reflectance factor:
- Fast varying
- Represents object

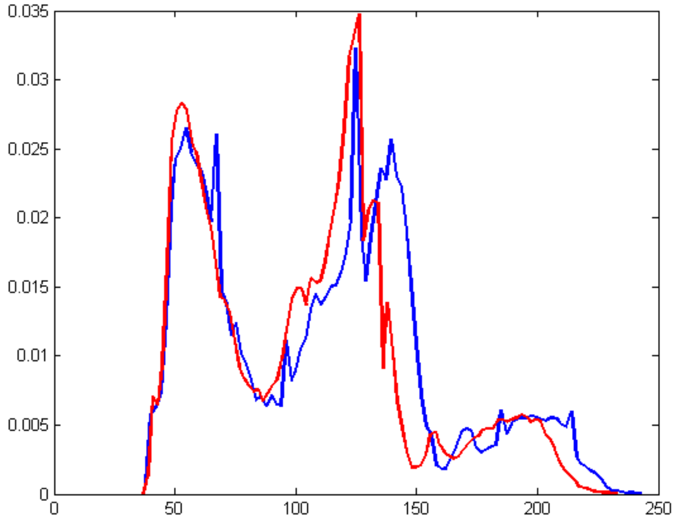
Illuminance factor:
- Slow varying
- Corresponds to illumination level

$$\alpha_{opt} = \underset{\alpha}{\arg\min}\ KL(h_l, h_r)$$

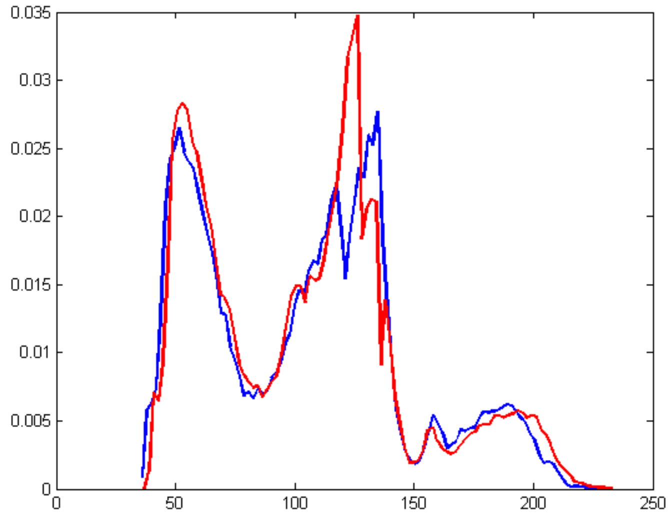
$$KL(h_l, h_r) = \frac{1}{2}\left[\sum h_l \log_2\left(\frac{h_l}{h_r}\right) + \sum h_r \log_2\left(\frac{h_r}{h_l}\right)\right]$$

Histograms of left and right images

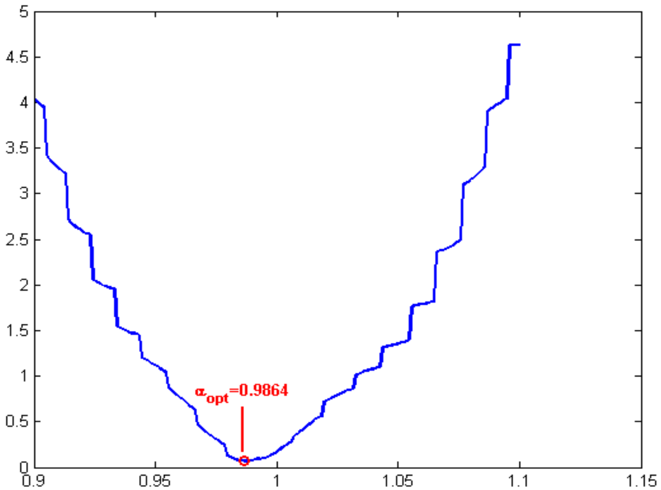
Before



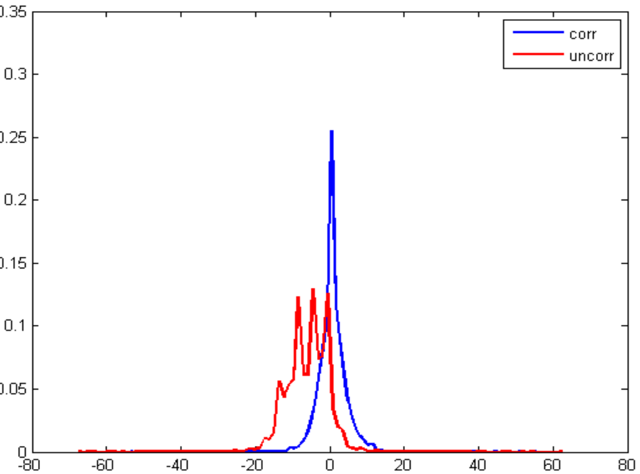
After



KL distance vs. Alpha



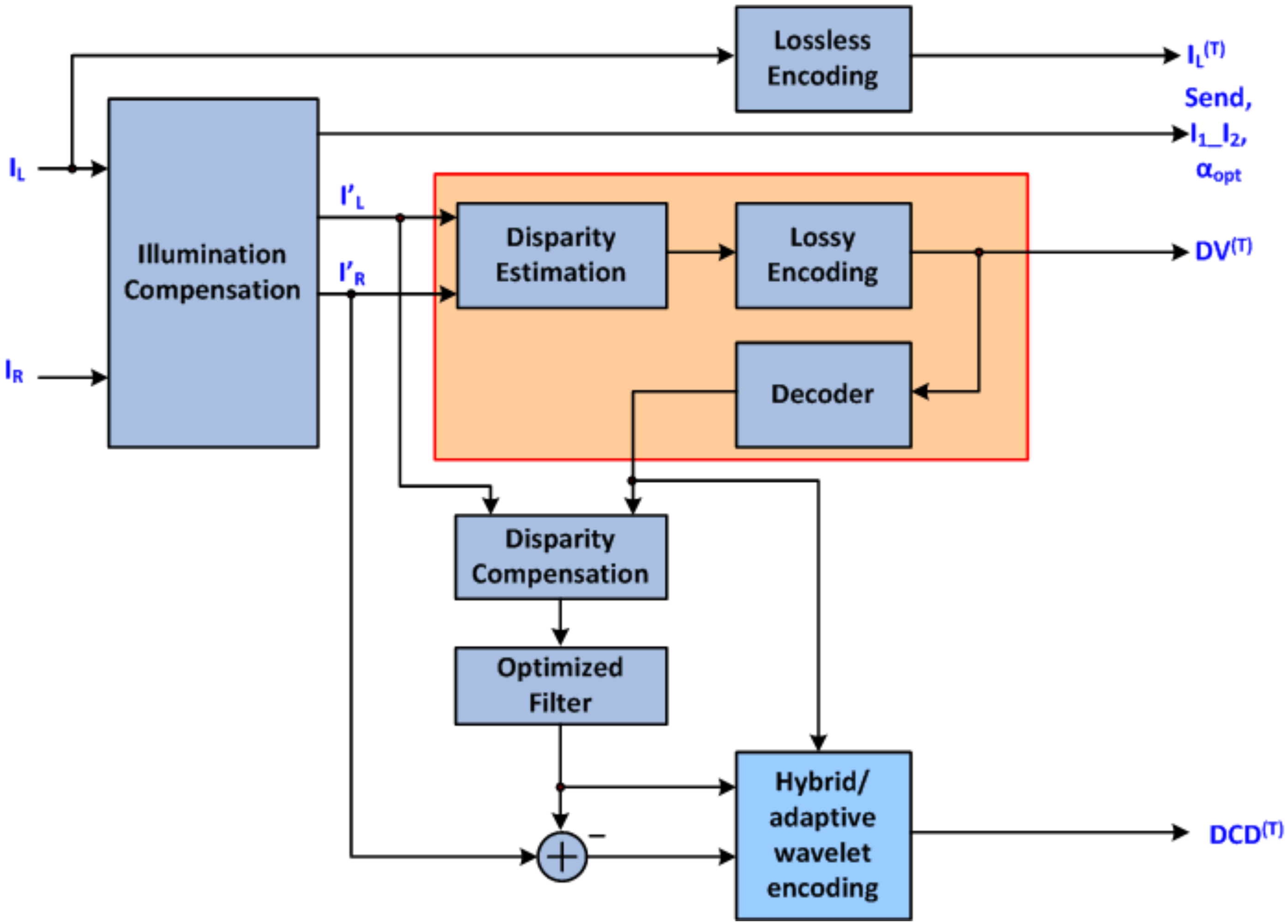
Histograms of residues



Results

Image	Before Compensation			After Compensation			Flags		α_{opt}
	Bitrate		KL Distance	Bitrate		KL Distance	$I_1\_I_2$	Send	
	DV $R^{(v)}$	Residue $R^{(e)}$		DV $R^{(v)}$	Residue $R^{(e)}$				
Apple	0.85	4.54	6.09	0.87	4.54	5.35	1	0	0.9997
Shrub	0.13	3.45	1.12	0.10	3.27	0.07	1	1	0.9867
J1	1.28	4.56	0.62	0.89	4.24	0.38	0	1	0.9857
Walker1	0.95	4.98	0.30	0.91	4.88	0.20	0	1	0.9692
Walker9	0.97	5.06	0.27	0.95	4.94	0.19	0	1	0.9687
Fruit	0.54	4.08	1.31	0.49	4.05	0.10	1	1	0.9892
Tsukuba	0.35	3.51	0.44	0.35	3.51	0.11	0	0	1.0003
HouseOf	0.72	5.37	1.09	0.71	5.36	0.01	0	1	0.9952

Disparity estimation and optimal quantization



Disparity estimation and optimal quantization (Contd...)

- Optical flow based disparity estimation Zach07

$$u^* = \underset{u}{\arg\min}\left[\int |\nabla u| dx + \lambda \int \left| I^{(l)}(x+u,y) - I^{(r)}(x,y) \right| dx \right]$$

  - Regularization: TV-norm
  - Data term: L_1-norm
- Choice of λ
  - Choose λ that minimizes total bit-rate

$$\underline{u}(\underline{x}) \equiv [u_x(\underline{x})\ u_y(\underline{x})]^T$$

$$\underline{u}^* = \underset{\underline{u}}{\arg\min}\ J(\underline{u})$$

$$J(\underline{u}) = \int |\nabla u| d\underline{x} + \lambda \int \left| I^{(l)}(\underline{x} + \underline{u}(\underline{x})) - I^{(r)}(\underline{x}) \right| d\underline{x}$$

$$u_y = 0 \Rightarrow \underline{u} \equiv u$$

$$J(u) \approx \int |\nabla u| d\underline{x} + \lambda \int \left| I^{(l)}(\underline{x}) + u(\underline{x}) I_x^{(l)}(\underline{x}) - I^{(r)}(\underline{x}) \right| d\underline{x}$$

$$J(u) \approx \int |\nabla u| d\underline{x} + \lambda \int |\rho(u,\underline{x})| d\underline{x}$$

$$J(u,\theta) = \int \left[\lambda |\rho(u)| + \frac{1}{2\theta}(u - v)^2 + |\nabla u| \right] d\underline{x}$$

For fixed v, solve for u

$$\underset{u}{\arg\min} \int \left[|\nabla u| + \frac{1}{2\theta}(u-v)^2 \right] dx$$

$$u = v - \theta\, div(\underline{p})$$

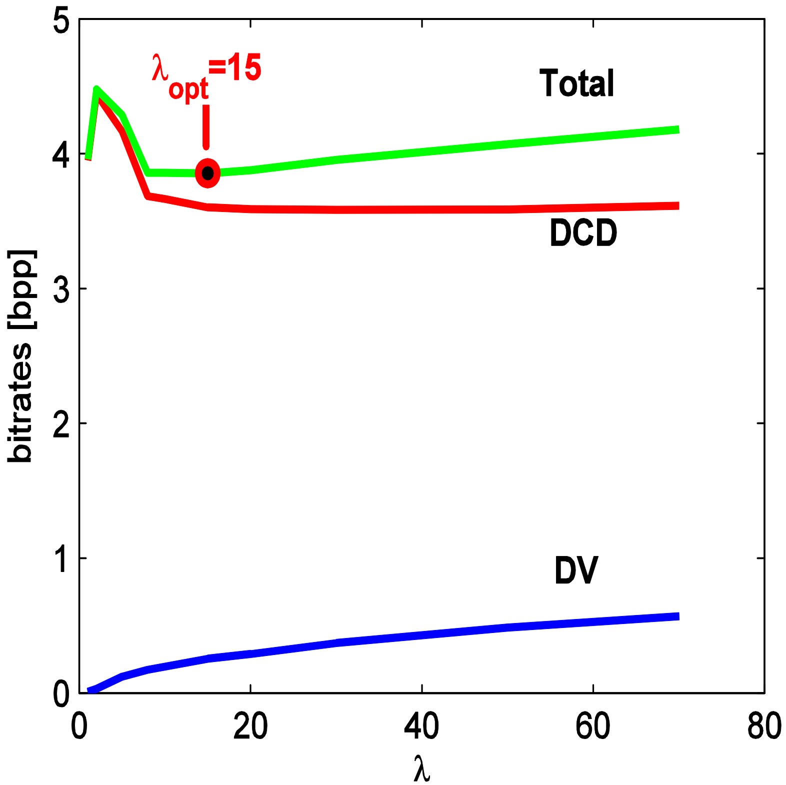
$$\underline{p}^{k+1} = \frac{\underline{p}^k + \tau\nabla(div(\underline{p}^k) - v/\theta)}{1 + |\tau\nabla(div(\underline{p}^k) - v/\theta)|} \qquad \underline{p}^0 = \underline{0}, \tau \le 1/8$$

For fixed u, solve for v

$$\underset{v}{\arg\min} \int \left[\lambda|\rho(u)| + \frac{1}{2\theta}(u-v)^2 \right] dx$$

$$v = u + \begin{cases} \lambda\theta I_x^{(l)} & if \;\; \rho(u) \le -\lambda\theta I^2{}_x^{(l)} \\ -\lambda\theta I_x^{(l)} & if \;\; \rho(u) > \lambda\theta I^2{}_x^{(l)} \\ -\dfrac{\rho(u)}{I_x^{(l)}} & otherwise \end{cases}$$

Choice of λ



Higher values of λ

•More weight on data term

•High bitrate for disparity but low for residue.

Lower values of λ

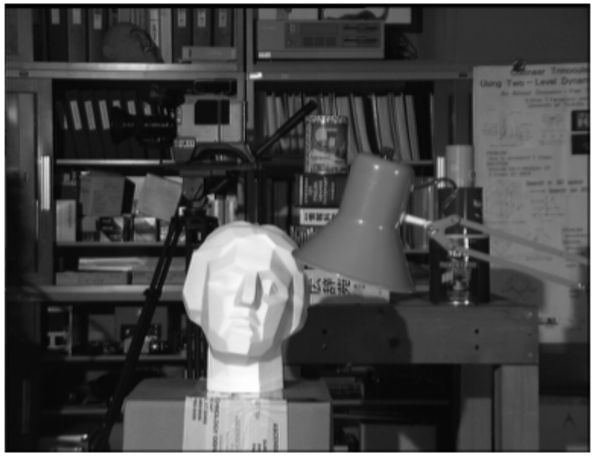
•Less weight on data term

•Low bitrate for disparity but high bitrate for residue

$$u^* = \underset{u}{\arg\min}\left[\int |\nabla u| dx + \lambda \int \left| I^{(l)}(x+u,y) - I^{(r)}(x,y) \right| dx\right]$$

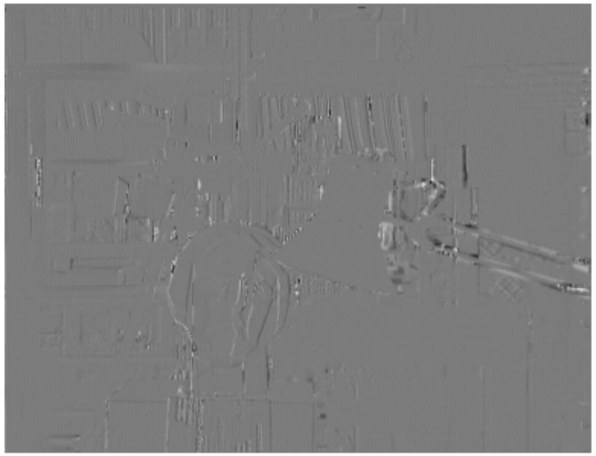
Results

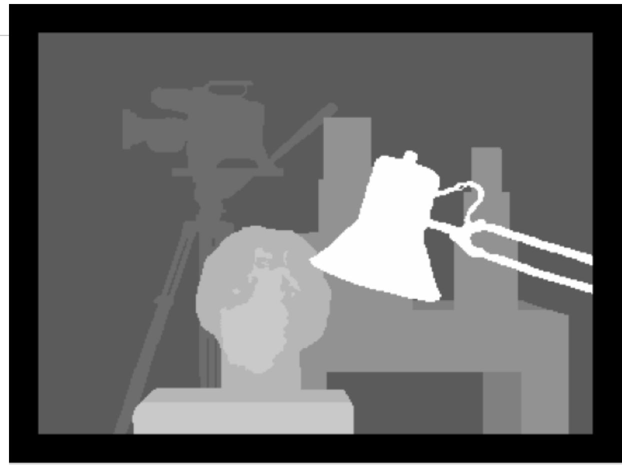
Image (Left only)



Estimated disparity map

Residual image

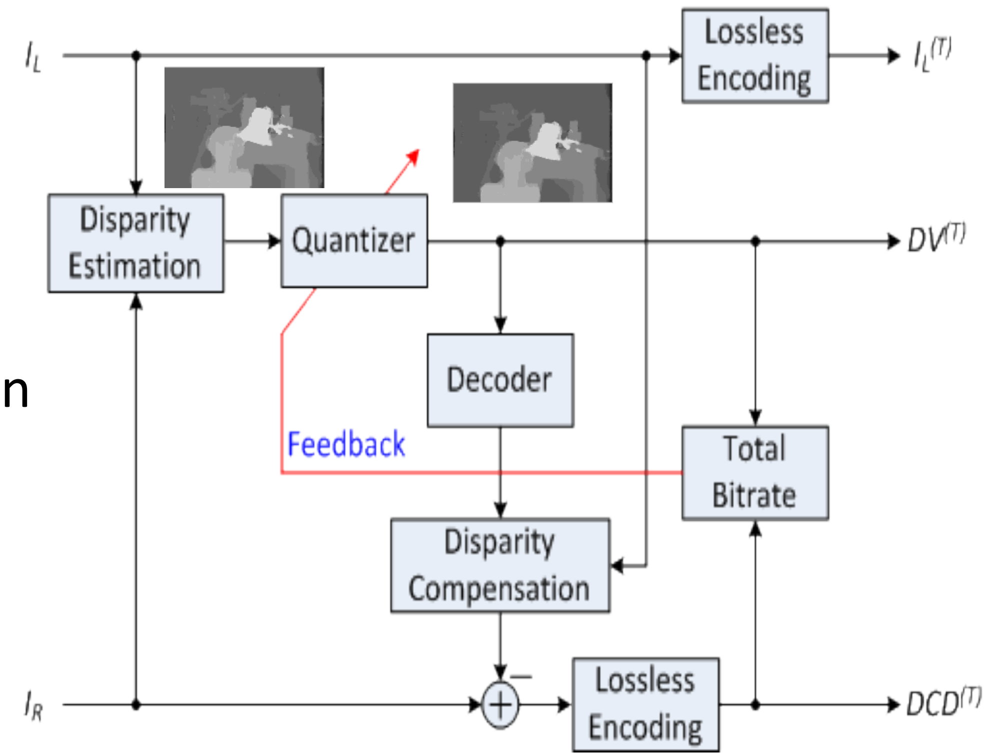




Ground truth

Optimal quantization of dense disparity map

- Allows tradeoff between bitrate of disparity map and that of residual image.
- Optimal quantization is chosen that minimizes the overall bitrate.
- Not possible with classical methods.
- No side information required.
- Compatible with JPEG2000.



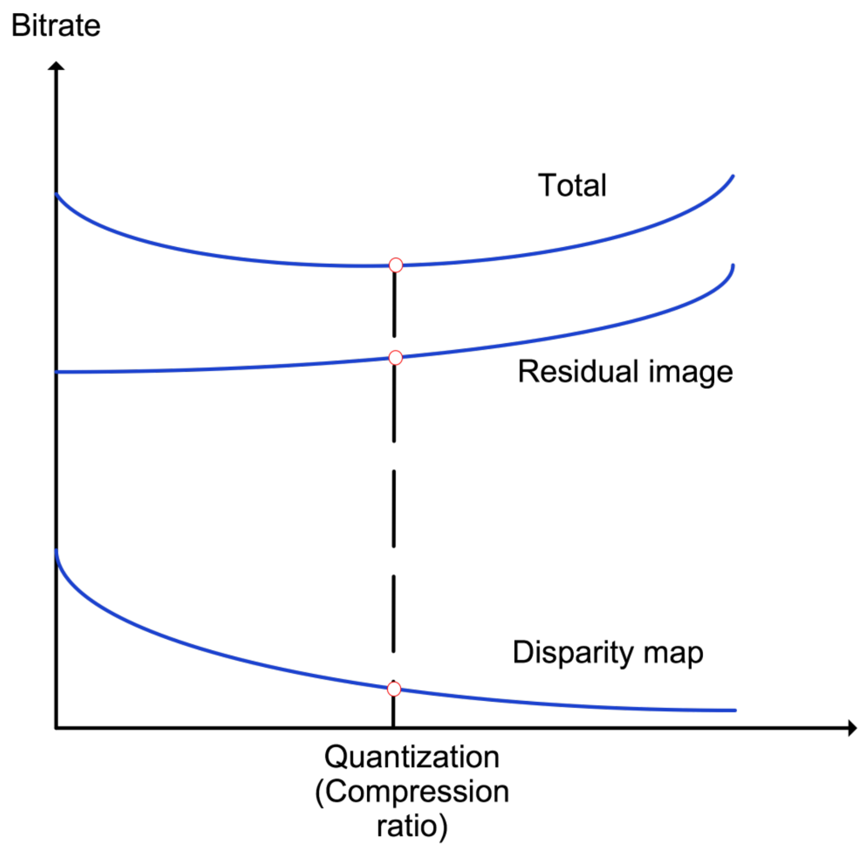
Amit Kumar K.C., Rony Darazi and Benoit Macq *'Optimal optical flow based disparity map estimation for lossless stereo image coding,' Electronic Imaging, IS&T/SPIE, San Francisco, California, USA, 2011*

Optimal quantization

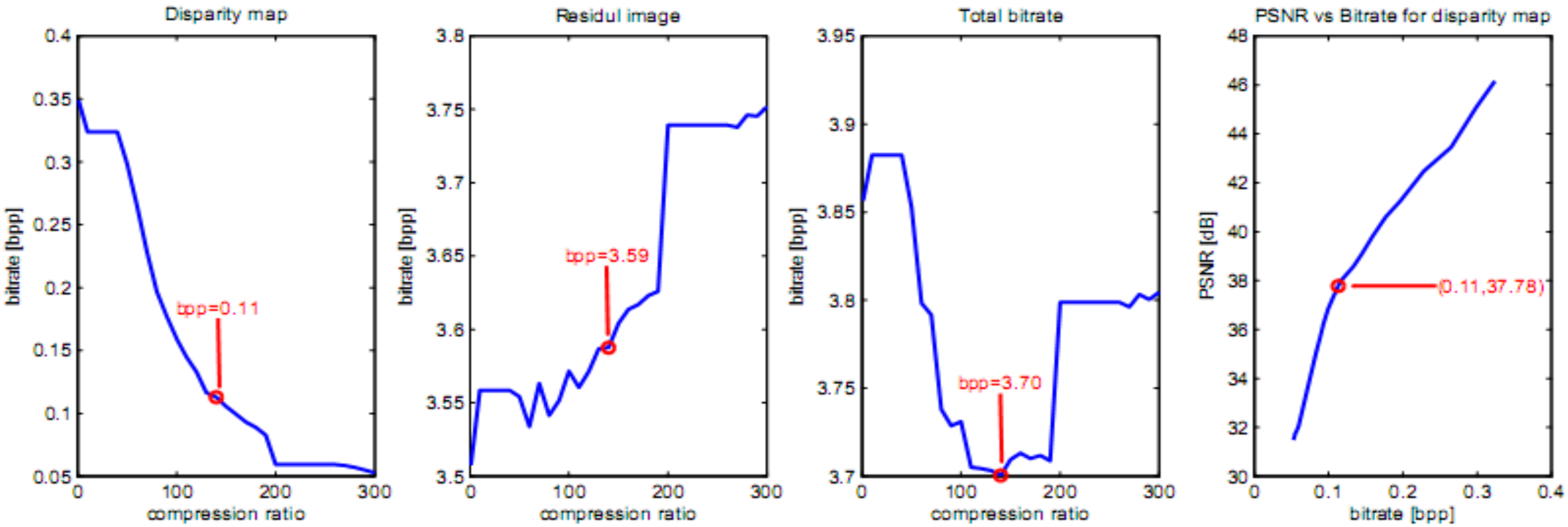
- Rates at which bitrates for disparity map and residual image vary are different.

- Initially, $R^{(v)}$ decreases rapidly, $R^{(e)}$ increases slowly.

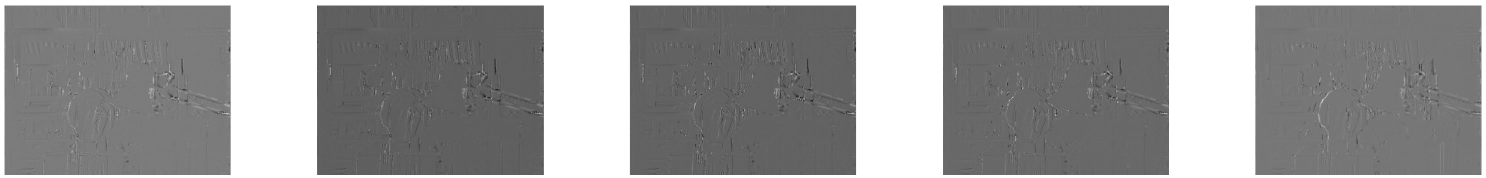
- After some quantization level, $R^{(v)}$ decreases slowly whereas $R^{(e)}$ increases substantially.



Results



comp. ratio=1 bpp= 0.35 psnr= Inf	comp. ratio=10 bpp= 0.32 psnr= 46.16	comp. ratio=50 bpp= 0.30 psnr= 45.08	comp. ratio=100 bpp= 0.16 psnr= 39.83	comp. ratio=200 bpp= 0.06 psnr= 32.04



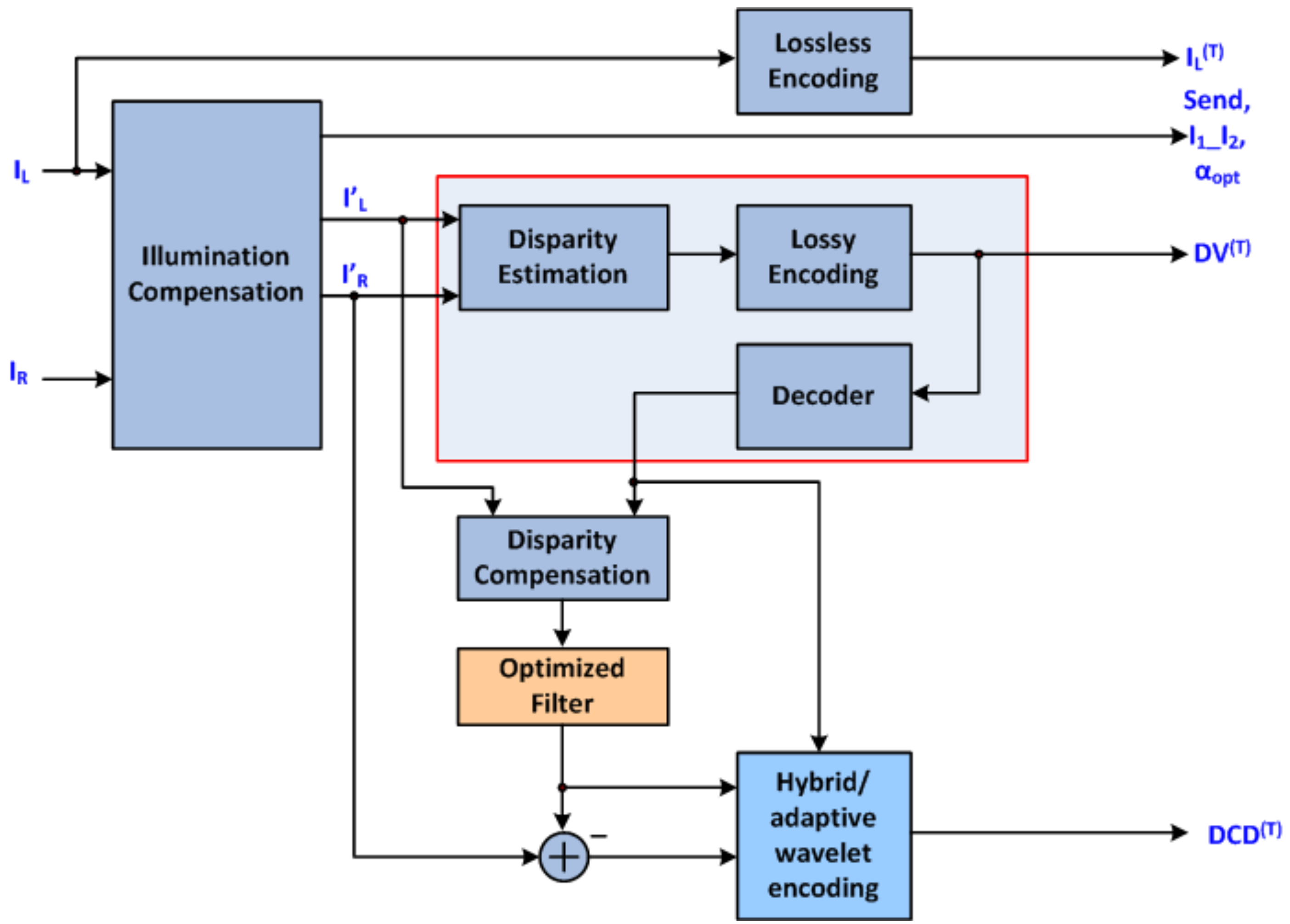
comp. ratio=1 bpp= 3.51	comp. ratio=10 bpp= 3.56	comp. ratio=50 bpp= 3.55	comp. ratio=100 bpp= 3.57	comp. ratio=200 bpp= 3.74

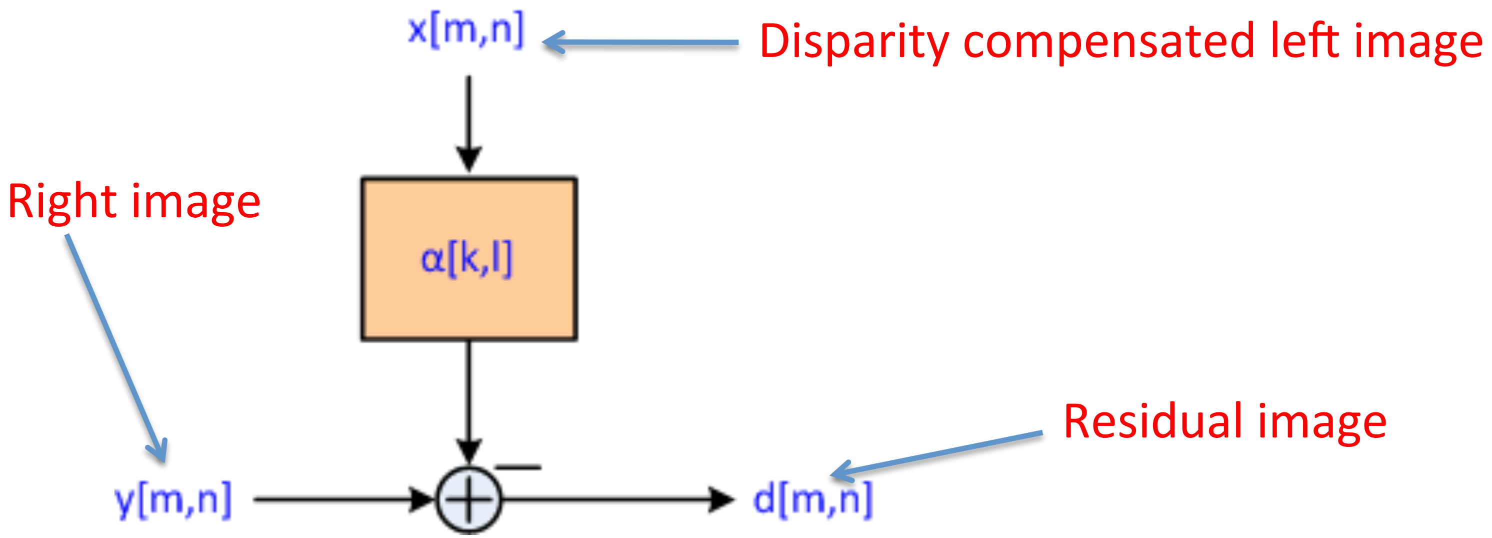
Results

Image	Bitrates [bpp] Before Optimization			Bitrates [bpp] After Optimization		
	Disparity	Residue	Total	Disparity	Residue	Total
Tsukuba	0.35	3.51	3.86	0.11	3.59	3.70
Fruit	0.49	4.05	4.54	0.06	4.10	4.16
Pentagon	0.64	5.11	5.75	0.06	5.21	5.27
Apple	0.87	4.54	5.41	0.10	4.72	4.82
HouseOf	0.71	5.36	6.07	0.08	5.45	5.53
Corridor	0.28	2.05	2.33	0.06	2.09	2.15
Birch	0.97	4.56	5.53	0.05	4.72	4.77
Shrub	0.10	3.27	3.37	0.05	3.27	3.32
Average	0.55	4.06	4.61	0.07	4.14	4.21

0.4 bpp

Optimized filter





- Minimum variance ➔ Minimum entropy
  [*For residual signal*]

- Design a filter <u>α</u> such that the variance of residual image is minimized.

Yule Walker Equations

$$\underline{\alpha}_{opt} = \underset{\underline{\alpha}}{\arg\min}\left[f(\underline{\alpha}) + \lambda g(\underline{\alpha})\right]$$

Gouze04, Platt88

$$f(\underline{\alpha}) = \sigma_d^{\ 2}$$

$$g(\underline{\alpha}) = 1 - \sum_{k,l} \alpha_{k,l}$$

$$E_{penalty} = \frac{c}{2} g^2(\underline{\alpha})$$

$$\dot{\alpha}_{p,q} = -\frac{\partial f(\underline{\alpha})}{\partial \alpha_{p,q}} - \lambda \frac{\partial g(\underline{\alpha})}{\partial \alpha_{p,q}} - c g(\underline{\alpha}) \frac{\partial g(\underline{\alpha})}{\partial \alpha_{p,q}}$$

$$\dot{\lambda} = +g(\underline{\alpha})$$

$$\frac{\partial f(\underline{\alpha})}{\partial \alpha_{p,q}} = -2\gamma_{XY}[p,q] + 2\sum_{k,l} \alpha_{k,l} \gamma_{XX}[p-k, q-l] + 2\mu_X \left(\mu_Y - \mu_X \sum_{k,l} \alpha_{k,l} \right)$$

$$\frac{\partial g(\underline{\alpha})}{\partial \alpha_{p,q}} = -1$$

$$\underline{X}_n = [x_1 \ldots x_K]^T \quad n = 1 \ldots MN \quad \rightarrow \textit{neighbourhood of current pixel at } n$$

$$\underline{X} = [\underline{X}_1^T \ldots \underline{X}_{MN}^T]^T \quad \rightarrow \textit{all pixels of input image}$$

$$\underline{Y} = [y_1 \ldots y_{MN}]^T \quad \rightarrow \textit{all pixels of output image}$$

$$\underline{\alpha} = [\alpha_1 \ldots \alpha_K]^T \quad \rightarrow \textit{filter coefficients}$$

$$\underline{\alpha}^* = \underset{\underline{\alpha}}{\arg\min} \frac{1}{MN} \left\| \underline{Y} - \underline{X}\,\underline{\alpha} \right\|_2^2 \quad \textit{such that } \underline{\mu}_X^T \underline{\alpha} = \mu_Y$$

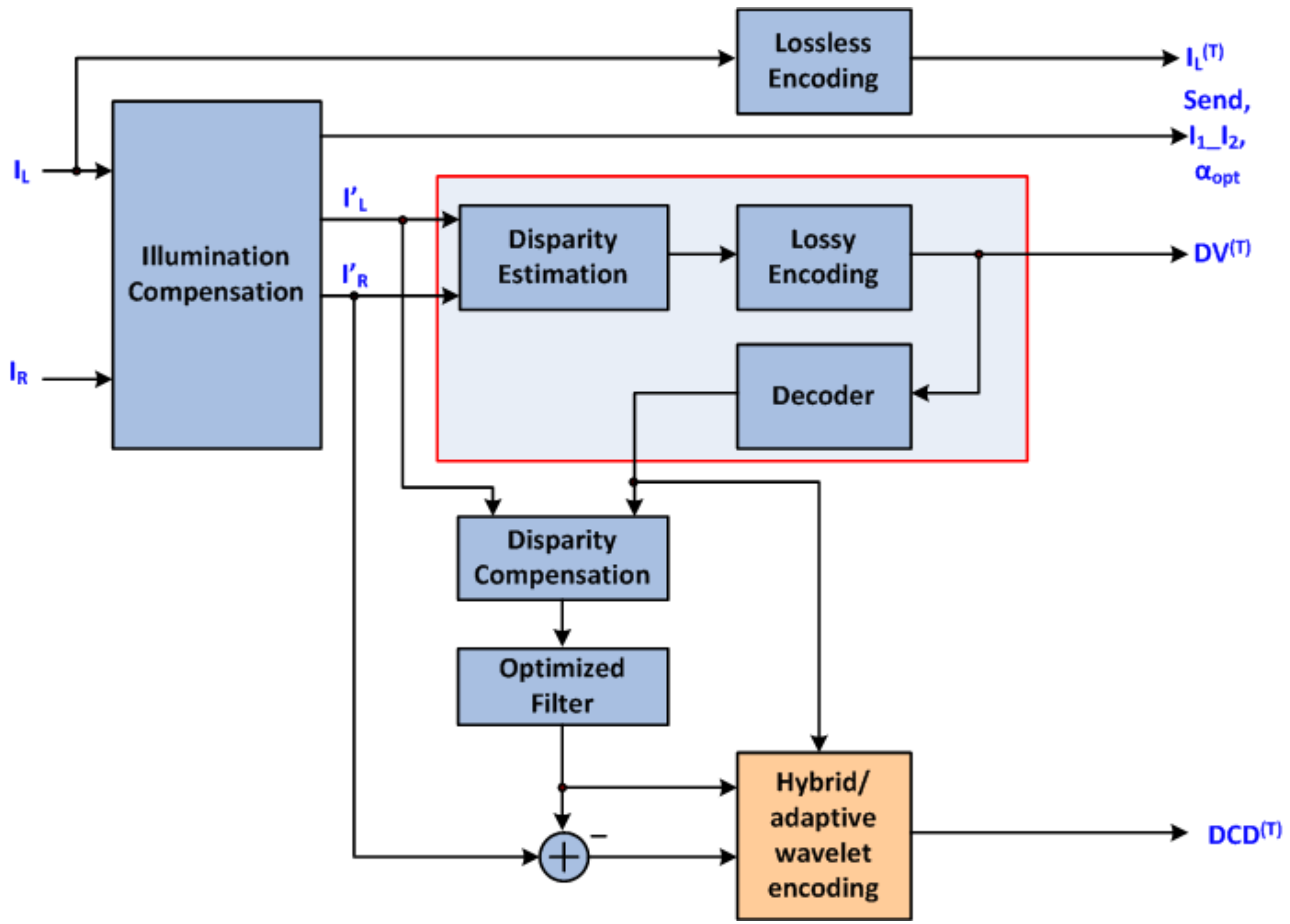
Minimum variance

Results

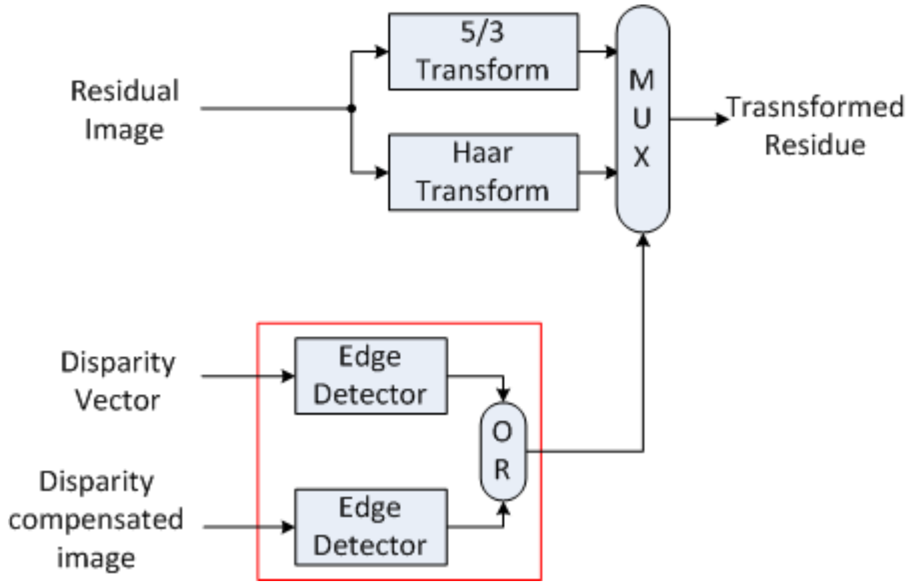
~0.5 bpp

Image	Filter size	Bitrate [bpp]		
		Unfiltered	Linear Filtering	Nonlinear Filtering
Fruit	3×3	4.15	3.55	3.59
	5×5		3.50	3.53
Pentagon	3×3	5.33	4.74	4.73
	5×5		4.79	4.77
Apple	3×3	4.75	4.40	4.39
	5×5		4.47	4.46
HouseOf	3×3	5.50	5.13	5.10
	5×5		5.16	5.13
Birch	3×3	4.70	4.27	4.30
	5×5		4.19	4.23

Hybrid wavelet encoding



Hybrid wavelet encoding (Contd…)



Ismael08

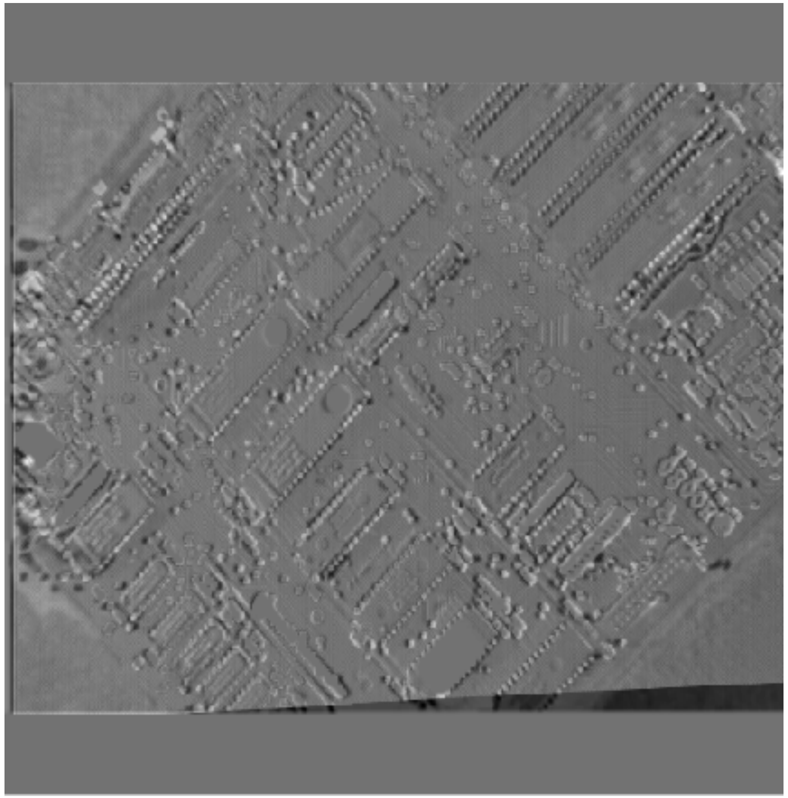
- Shorter filters around edges of the residue
  - Due to occlusions and mismatches
  - Reduces ringing artifacts
  - Haar wavelet transform
- Longer filters for smooth areas
  - 5/3 wavelet transform

Approximate the edge map by

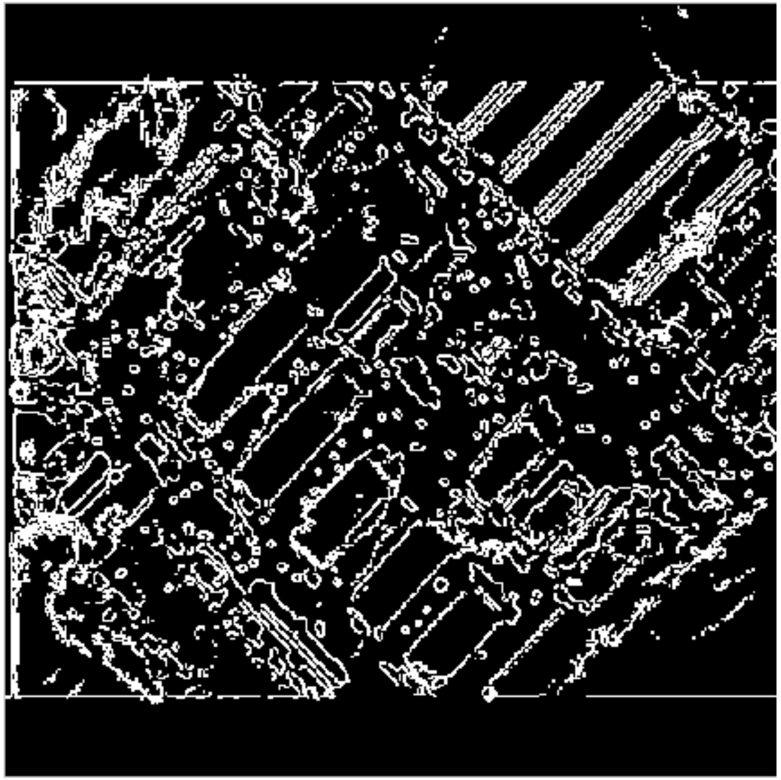
•Edges in disparity map→characterize occlusion

•Edges in disparity compensated left image → characterize mismatch

Results



Approximated edge



Residue

*Rony Darazi, **Amit Kumar K.C**. and Benoit Macq, "Using Depth Map for Directional Adaptive Lifting Scheme for Stereo Image Residuals", Submitted for ISCAS11*

Results (Contd…)

~0.1 bpp

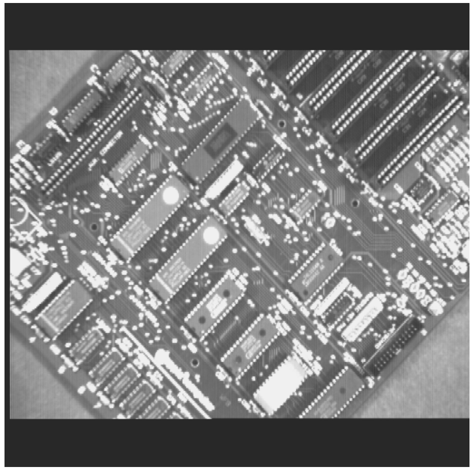
Image	Entropy [bpp]		
	5/3 Transform	Haar Transform	Hybrid Transform
Fruit	2.44	2.64	2.38
Pentagon	4.31	4.29	4.20
Baseball	4.71	4.86	4.52
Tsukuba	2.49	2.36	2.31
Ball1	2.94	3.33	2.89
Apple	4.09	3.98	3.91
HouseOf	4.22	4.28	4.18
Corridor	1.79	1.79	1.54
Pm	3.02	3.12	2.96
Book	2.95	3.11	2.80

Putting it all together

Image	Baseline Scheme $I^{(l)}, I^{(e)}$	Vector Lifting Scheme *Kaaniche et al.*		Proposed Scheme		
		VLS-I	VLS-II	J=2	J=3	J=4
Fruit	4.05	3.78	3.72	3.83	3.73	3.71
Shrub	3.73	3.81	3.63	3.53	3.45	3.44
Birch	4.52	4.44	4.37	4.19	4.01	3.96
Pentagon	5.37	5.12	5.04	5.10	5.00	4.99
Average	4.42	4.29	4.19	4.18	4.08	4.03

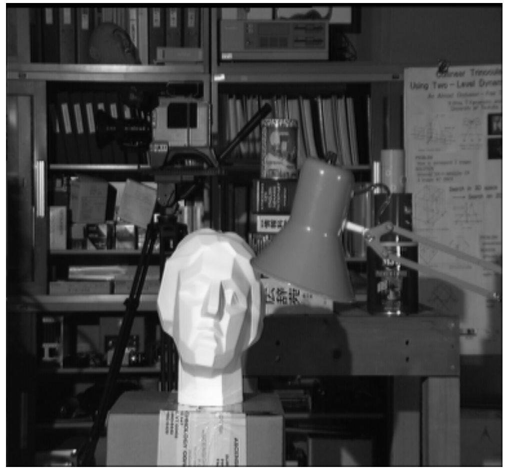
~0.1 bpp
Scalable

Test Images



Apple

Fruit



Tsukuba

HouseOf

Pentagon

Shrub

Bibliography

- Woontack Woo and Antonio Ortega, “Stereo image compression with disparity compen- sation using the mrf model,” in Proc. SPIE VCIP, 1996, pp. 28–41.

- M. Kaaniche, A. Benazza-Benyahia, B. Pesquet-Popescu, and J.-C. Pesquet, “Vector lifting schemes for stereo image coding,” Image Processing, IEEE Transactions on, vol. 18, no. 11, pp. 2463 –2475, nov. 2009.

- A. Gouze, M Antonini, M Barlaud, and B. Macq, “Design of signal adapted multidimen- sional lifting scheme for lossy coding,” 2004, pp. 1589–1603.

- C. Zach, T. Pock, and H. Bischof, “A duality based approach for realtime tv-l1 optical flow,” in Pattern Recognition (Proc. DAGM), Heidelberg, Germany, 2007, pp. 214–223.

Thank you very much