

WAVELET INPAINTING DRIVEN IMAGE COMPRESSION VIA COLLABORATIVE SPARSITY AT LOW BIT RATES

Chen Zhao¹, Jian Zhang², Siwei Ma¹, Wen Gao¹

¹Institute of Digital Media, Peking University, Beijing, China

²School of Computer Science and Technology, Harbin Institute of Technology, Harbin, China

ABSTRACT

To overcome the unsatisfactory encoding quality of conventional image compression methods at low bit rates, the idea of downsampling prior to encoding and upsampling after decoding turns out to be a good solution. Based on this paradigm, we propose a low-bit-rate image compression algorithm by use of the novel wavelet inpainting technique via collaborative sparsity. Superior to the existing methods which operate the sampling in the space domain, we merge the wavelet transform in the downsampling stage, which is verified to be able to preserve much more information. By investigating the local two-dimensional sparsity and the nonlocal three-dimensional sparsity of the image simultaneously, a collaborative sparsity model is exploited to restore the full-resolution image from the decoded downsampled image. Finally a Split Bregman based iterative algorithm is developed to solve the optimization problem. Experimental results demonstrate obvious visual quality improvements, as well as PSNR gains, compared to the state-of-the-art methods under various low bit rates.

Index Terms— Image compression, low bit rates, wavelet inpainting, sparsity

1. INTRODUCTION

Conventional image compression methods [1], e.g. JPEG and JPEG 2000, seek to encode every single pixel in the original image. This may well preserve the original information when bandwidth permits. In the case of low bit rate transmission such as the mobile network, however, severe blockiness and other coding artifacts would arise as a result [2, 3]. Each pixel could be merely allocated limited number of bits on average, and the large quantization step size adopted would greatly suppress the information left in the reconstructed picture. Consequently, severe deterioration of image compression quality would occur.

Knowing that images tend to be generated by oversampling [4], a lot of redundancy exists and thus some representations of the pixels could be removed. According to research discoveries, downsampling prior to encoding and upsampling after decoding can improve the quality of coded

image at low bit rates [5-7]. For this reason, some works are focusing on the downsampling and upsampling stage to derive a better encoding performance.

W. Lin et al [8] proposes an adaptive downsampling method, which determines the downsampling ratio/direction and quantization step for each macroblock based upon the local visual significance of the signal. It outperforms the JPEG coding method to some extent but pales in comparison to the more advanced JPEG 2000 method. Besides, its modification to the coding framework plus the macroblock-based feature makes it incompatible with traditional coding standards. An improved scheme called the collaborative adaptive down-sampling and upconversion (CADU) is proposed by X. Wu et al [9]. It filters the image in a spatially varying and directional way ahead of the uniform downsampling in the image space and encodes it using a third party codec. At the decoder, the low-resolution image is upconverted to the original resolution in a constrained least squares restoration process. This method can be applied without change to the current image coding standard, and achieves better results than the one in [8]. Yet the restored image quality is restrained by the space-domain downsampling, which cannot preserve adequate information for the image up conversion.

In this paper, we propose a wavelet inpainting driven image compression algorithm (WIDIC) at low bit rates. Our contribution is threefold. First, we put forward the image compression framework of downsampling and upsampling in the wavelet domain, which can be combined with any image codec. The low frequency component in wavelet domain is believed to retain more information than the simply downsampled one in the space domain. Second, we formulate the process of restoring the high-resolution image from the decoded low-resolution one as a wavelet inpainting optimization problem. We adopt a collaborative sparsity model as the regularization term, which adaptively enforces local two-dimensional smoothness and nonlocal three-dimensional self-similarity simultaneously in a hybrid space-transform domain. Thus it substantially exploits the intrinsic features of the natural image and can offer high quality results. Third, a Split-Bregman based iterative algorithm is developed to efficiently solve the optimization problem. This proposal

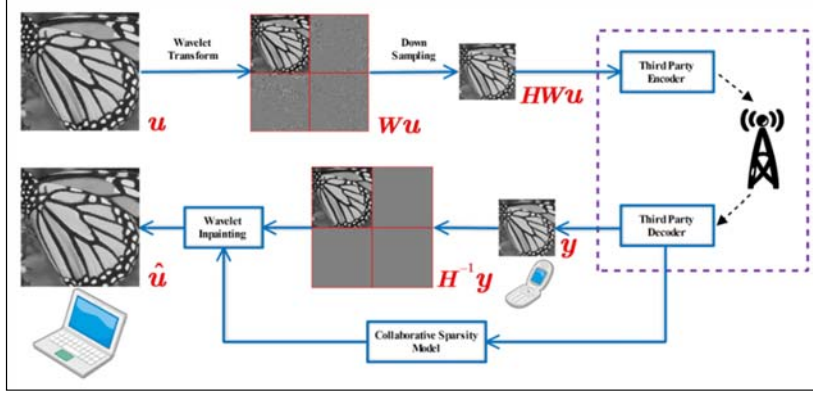


Fig. 1: Block diagram of the proposed wavelet inpainting driven image compression system.

provides a novel idea of integrating the prior model of an image into the compression framework. It is compatible with any third-party compression techniques and outperforms the highly evaluated JPEG 2000 and the above-mentioned CADU.

The rest of the paper is organized as follows. Section II describes the details of the proposed wavelet inpainting driven image compression algorithm. Section III presents the experimental results. Lastly, Section IV contains our summary and derived conclusions.

2. PROPOSED WIDIC ALGORITHM

2.1. Problem Formulation

As shown in Fig. 1, the image first goes through the wavelet transform and then is downsampled into a low-resolution image in the transform domain. Encoded by any coding technique, the downsampled image is to be transmitted through the network and received by the other side. The entire compression process can be expressed as

$$\mathbf{y} = \mathbf{H}\mathbf{W}\mathbf{u} + \mathbf{n} \quad (1)$$

where \mathbf{u} represents the original high-resolution image and \mathbf{y} is the directly reconstructed low-resolution image from the decoder. \mathbf{W} denotes the wavelet transform operator and \mathbf{H} is the downsampling operator, the two of which jointly downsample the original image into a low-resolution one in the wavelet transform domain. \mathbf{n} is the noise induced during the encoding and transmission periods, which is assumed to be the additive Gaussian white noise.

After decoding, the reconstructed low-resolution image \mathbf{y} can be displayed on devices with small screens such as the mobile phone. For regular-sized screens like the computer, it is expected to be upsampled to a high-resolution image in the wavelet domain, which may be considered as an inverse process of the downsampling. So in order to find a restored image $\hat{\mathbf{u}}$ closest to the original image \mathbf{u} , we can formulate it into an optimization problem as follows.

$$\hat{\mathbf{u}} = \underset{\mathbf{u}}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{H}\mathbf{W}\mathbf{u} - \mathbf{y}\|_2^2 \quad (2)$$

Since Eq. (2) above is an ill-posed problem and no specific solution could be found, we need to incorporate a regularization term and hereby get

$$\hat{\mathbf{u}} = \underset{\mathbf{u}}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{H}\mathbf{W}\mathbf{u} - \mathbf{y}\|_2^2 + \alpha \cdot \Psi(\mathbf{u}) \quad (3)$$

where $\Psi(\mathbf{u})$ stands for the regularization term, which describes the prior knowledge of the image. α is a regularization parameter that controls the tradeoff between the two terms.

Recently, many works concentrating on utilization of both local smoothness and nonlocal self-similarity have achieved great success in image restoration applications [10, 14 and 15]. Therefore, to find a suitable regularization term to well exhibit the intrinsic characteristics of the image, this paper adopts the collaborative sparsity model, which is first proposed in [10, 14] for image compressive sensing recovery and written as

$$\Psi(\mathbf{u}) = \tau \cdot \Psi_{L2D}(\mathbf{u}) + \lambda \cdot \Psi_{N3D}(\mathbf{u}) \quad (4)$$

In this model, two sparse features of the image are investigated, the local smoothness and the nonlocal self-similarity, which are denoted by $\Psi_{L2D}(\mathbf{u})$ and $\Psi_{N3D}(\mathbf{u})$ respectively. τ and λ are their corresponding parameters and collaboratively weigh the contributions of the two terms.

Specifically, the local smoothness is formulated as two-dimensional sparsity in the space domain, which can be described by the follow equation.

$$\Psi_{L2D}(\mathbf{u}) = \|\mathcal{D}\mathbf{u}\|_1 = \|\mathcal{D}_v\mathbf{u}\|_1 + \|\mathcal{D}_h\mathbf{u}\|_1 \quad (5)$$

where $\mathcal{D} = [\mathcal{D}_v; \mathcal{D}_h]$ and are vertical and horizontal finite gradient operators. Given the observation that image gradient values are close to zero and roughly conform to the Laplacian Distribution, we suppress the L1-norm of the gradients to ensure the image local smoothness.

The nonlocal self-similarity is exploited as three-dimensional sparsity in the transform domain. We first stack

similar image patches in a three dimensional way, and then transform the group of patches into a domain where the coefficients demonstrate salient sparsity. This feature could be characterized by

$$\Psi_{N3D}(\mathbf{u}) = \sum_{k=1} \|\mathbf{T}^{3D}(\mathbf{Z}_{u_k})\|_0 \quad (6)$$

where \mathbf{T}^{3D} the operator of a three-dimensional transform and \mathbf{Z}_{u_k} is a three-dimensional array containing similar patches corresponding to the image patch \mathbf{u}_k .

Readers may refer to [10] for more details about the formulation of the collaborative sparsity model. Moreover, it is worth emphasizing that any image prior model can be incorporated into our proposed WIDIC algorithm. For this reason, it establishes an interesting and direct connection between natural image priors and image compression.

So far, Eq. (3) could be rewritten as

$$\arg\min_{\mathbf{u}} \frac{1}{2} \|\mathbf{H}\mathbf{W}\mathbf{u} - \mathbf{y}\|_2^2 + \tau \cdot \Psi_{L2D}(\mathbf{u}) + \lambda \cdot \Psi_{N3D}(\mathbf{u}) \quad (7)$$

To solve the optimization problem defined in Eq. (7), a Split Bregman based iterative scheme is devised and will be elaborated in Subsection 2.2.

2.2. Split Bregman based Iterative Algorithm

First, for convenient representation we introduce the proximal map $\text{prox}_t(g)(\mathbf{x})$ defined by [11].

$$\text{prox}_t(g)(\mathbf{x}) = \arg\min_{\mathbf{u}} \left\{ \frac{1}{2} \|\mathbf{u} - \mathbf{x}\|_2^2 + t \cdot g(\mathbf{u}) \right\} \quad (8)$$

In Eq (8), g refers to a proper closed convex function and t is a positive scalar.

Note that Eq. (7) is essentially not a convex problem and quite difficult to solve directly. Applying Split Bregman algorithm [12] to Eq. (7) leads to the following iterative steps:

$$(\hat{\mathbf{u}}^{(j+1)}, \hat{\mathbf{x}}^{(j+1)}, \hat{\mathbf{w}}^{(j+1)}) = \arg\min_{\mathbf{u}, \mathbf{w}, \mathbf{x}} \frac{1}{2} \|\mathbf{H}\mathbf{W}\mathbf{u} - \mathbf{y}\|_2^2 + \tau \cdot \Psi_{L2D}(\mathbf{w}) \quad (9)$$

$$\begin{aligned} & + \lambda \cdot \Psi_{N3D}(\mathbf{x}) + \frac{\mu_1}{2} \|\mathbf{u} - \mathbf{w} - \mathbf{b}^{(j)}\|_2^2 + \frac{\mu_2}{2} \|\mathbf{u} - \mathbf{x} - \mathbf{c}^{(j)}\|_2^2 \\ & \mathbf{b}^{(j+1)} = \mathbf{b}^{(j)} - \left(\hat{\mathbf{u}}^{(j+1)} - \hat{\mathbf{w}}^{(j+1)} \right) \\ & \mathbf{c}^{(j+1)} = \mathbf{c}^{(j)} - \left(\hat{\mathbf{u}}^{(j+1)} - \hat{\mathbf{x}}^{(j+1)} \right) \end{aligned} \quad (10)$$

Further, an alternating direction technique is employed for the split problem in (9), which alternatively minimizes one variable while keeping the other variables. Thus, the following three sub-problems are yielded. Similar to [10], we argue that every separated sub-problem has an efficient solution. For simplicity, the subscript j is omitted without confusion.

2.2.1. \mathbf{u} sub-problem

Given \mathbf{w} and \mathbf{x} , the optimization problem associated with \mathbf{u} can be expressed as

$$\mathbf{u} = \arg\min_{\mathbf{u}} \frac{1}{2} \|\mathbf{H}\mathbf{W}\mathbf{u} - \mathbf{y}\|_2^2 + \frac{\mu_1}{2} \|\mathbf{u} - \mathbf{w} - \mathbf{b}\|_2^2 + \frac{\mu_2}{2} \|\mathbf{u} - \mathbf{x} - \mathbf{c}\|_2^2 \quad (11)$$

Since (11) is a minimization problem of strictly convex quadratic function, there is a closed form expressed as

$$\hat{\mathbf{u}} = (\mathbf{B}^T \mathbf{B} + \mu \mathbf{I})^{-1} \cdot \mathbf{q} \quad (12)$$

where $\mathbf{B} = \mathbf{H}\mathbf{W}$, $\mathbf{q} = \mathbf{B}^T \mathbf{y} + \mu_1(\mathbf{b} + \mathbf{w}) + \mu_2(\mathbf{c} + \mathbf{x})$, \mathbf{I} is identity matrix and $\tilde{\mu} = \mu_1 + \mu_2$. To avoid the operation of matrix inversion, owing to the particular structure of Matrix \mathbf{B} that satisfies $\mathbf{B}^T \mathbf{B} = \mathbf{I}$, applying the Sherman-Morrison-Woodbury matrix inversion formula to (12) yields

$$\hat{\mathbf{u}} = \frac{1}{\tilde{\mu}} (\mathbf{I} - \frac{1}{1+\tilde{\mu}} \mathbf{W}^T \mathbf{H}^T \mathbf{H} \mathbf{W}) \cdot \mathbf{q} \quad (13)$$

2.2.2. \mathbf{w} sub-problem

With the aid of \mathbf{u}, \mathbf{x} , the \mathbf{w} sub-problem is equivalent to

$$\mathbf{w} = \text{prox}_{\gamma}(\Psi_{L2D})(\mathbf{p}) \quad (14)$$

where $\mathbf{p} = \mathbf{u} - \mathbf{b}$ and $\gamma = \tau / \mu_1$.

Indeed, the proximal map $\text{prox}_{\gamma}(\Psi_{L2D})(\mathbf{p})$ can be solved efficiently by fast iterative shrinkage/thresholding algorithm (FISTA) [11].

2.2.3. \mathbf{x} sub-problem

With the recovered \mathbf{u} and \mathbf{w} , \mathbf{x} sub-problem is

$$\hat{\mathbf{x}} = \text{prox}_{\alpha}(\Psi_{N3D})(\mathbf{r}) \quad (15)$$

where $\mathbf{r} = \mathbf{u} - \mathbf{c}$ and $\alpha = \lambda / \mu_2$.

According to [10], the solution to the proximal map $\text{prox}_{\alpha}(\Psi_{N3D})(\mathbf{r})$ can be obtained analytically by just utilizing hard thresholding function.

In light of all derivations above, the complete description of proposed WIDIC via collaborative sparsity model is given in Table 1.

Table 1: The Proposed Wavelet Inpainting Driven Image Compression via Collaborative Sparsity

Input: the received low resolution image \mathbf{y} and the sub-sampling matrix \mathbf{H} .
Initialization: $\hat{\mathbf{x}}^{(0)} = \hat{\mathbf{u}}^{(0)} = \hat{\mathbf{w}}^{(0)} = \mathbf{y}, \mathbf{b}^{(0)} = \mathbf{c}^{(0)} = \mathbf{0}, k = 0, \mu_1, \mu_2, \tau, \lambda;$
Loop $k = 1, 2, \dots$ Solve \mathbf{u} sub-problem to achieve $\hat{\mathbf{u}}$; Solve \mathbf{w} sub-problem to achieve $\hat{\mathbf{w}}$; Solve \mathbf{x} sub-problem to achieve $\hat{\mathbf{x}}$; Update \mathbf{b} and \mathbf{c} ; Until stopping criterion is satisfied
Output: Final restored image $\hat{\mathbf{x}}$.

3. EXPERIMENTAL RESULTS

To evaluate the proposed WIDIC method, we carry out the experiments on four groups of standard gray-level images: “Butterfly”, “Lena”, “Leaves” and “Boat”. Compressing these test images under various bit rates, we compare the restoration performance of our method with that of JPEG 2000 (J2K for short) and CADU [9]. Both subjective quality and objective quality results will be demonstrated.

Table 2: PSNR comparisons among different methods at various bit rates

Rate (bpp)	<i>Butterfly</i>			<i>Lena</i>			<i>Leaves</i>			<i>Boat</i>		
	J2K	CADU	WIDIC	J2K	CADU	WIDIC	J2K	CADU	WIDIC	J2K	CADU	WIDIC
0.10	18.91	18.75	19.17	25.38	25.08	25.69	17.85	18.20	18.07	23.54	23.35	23.66
0.15	20.82	20.84	20.76	27.16	27.16	27.69	19.77	19.75	20.11	25.11	24.95	25.53
0.20	21.99	22.17	22.43	28.53	28.38	28.87	21.12	21.19	21.70	26.29	25.93	26.54
0.25	23.18	23.24	23.88	29.82	29.83	30.17	22.62	22.52	23.09	27.10	26.74	27.42
0.30	24.01	24.11	25.06	30.73	30.75	31.33	23.26	23.48	24.34	27.97	27.18	27.92
0.35	24.85	25.17	25.97	31.16	31.37	31.74	24.29	24.62	25.72	28.80	27.62	28.37
0.40	25.44	25.86	26.30	32.23	31.87	32.61	24.95	24.98	26.39	29.71	27.80	28.78



Fig. 2: Comparison of different methods at 0.25 bpp. Left: J2K (PSNR=23.18 dB, SSIM=0.7699); Middle: CADU (PSNR=23.24 dB, SSIM=0.7748); Right: WIDIC (PSNR=23.88 dB, SSIM=0.8130).



Fig. 3: Comparison of different methods at 0.25 bpp. Left: J2K (PSNR=29.82 dB, SSIM=0.8297); Middle: CADU (PSNR=29.83 dB, SSIM=0.8250); Right: WIDIC (PSNR=30.17 dB, SSIM=0.8340).



Fig. 4: Comparison of different methods at 0.35 bpp. Left: J2K (PSNR=24.29 dB, SSIM=0.8298); Middle: CADU (PSNR=24.62 dB, SSIM=0.8394); Right: WIDIC (PSNR=25.72 dB, SSIM=0.8878).



Fig. 5: Comparison of different methods at 0.35 bpp. Left: J2K (PSNR=28.80 dB, SSIM=0.7906); Middle: CADU (PSNR=27.62 dB, SSIM=0.7847); Right: WIDIC (PSNR=28.37 dB, SSIM=0.7961).

The PSNR comparisons among the three methods are shown in Table 2. At each bit rate, the best PSNR results are marked in bold for each image. We can see that our WIDIC always beat JPEG 2000 and CADU for almost all the cases under extremely low bit rates from 0.10 bpp to 0.30 bpp. Even when the bit rate reaches as high as 0.3 bpp to 0.40 bpp, it can

still demonstrate obvious advantages over the other methods, and remain the best for most of the test groups. It can provide PSNR gains up to 1.44 dB compared to JPEG 2000, and up to 1.42 dB compared to WIDIC. In contrast, CADU can only achieve slight PSNR increases than JPEG 2000.

Fig. 2 through Fig. 5 give the visual quality comparisons of the restored images with the three methods. Fig. 2 and Fig. 3 are the restoration results of the images “Butterfly” and “Lena” under the bit rate 0.25 bpp, respectively. It is easy to notice that compared to JPEG 2000 and CADU, which both lead to annoying artifacts, our WIDIC can produce clear images which are much more pleasant to human eyes. Fig. 4 and Fig. 5 are restoration results of the images “Leaves” and “Boat” under the bit rate 0.35 bpp, respectively. At the higher bit rate, WIDIC yields the purest image of the three, which is consistent with the PSNR and SSIM [13] results in Fig. 4. In Fig. 5, although WIDIC cannot achieve a higher PSNR value than JPEG 2000, it does generate a better visual image without the burr effects surrounding the mast.

4. CONCLUSION

In this paper, we propose a wavelet inpainting driven image compression algorithm at low bit rates. Instead of directly encoding the original high-quality image, we base our scheme on the framework of downsampling prior to encoding and upsampling after decoding. We believe that downsampling in the wavelet domain can better preserve the original information than the space domain through sufficient experiments. Hence, we integrate the wavelet inpainting technique at the receiver side and devise a joint statistic model by exploiting the local and nonlocal sparsity of the image. Extensive experiments demonstrate that our method can distinctly outperform the well-known coding standard JPEG 2000 and the most state-of-the-art method CADU, both in subjective image quality and PSNR.

5. ACKNOWLEDGEMENTS

This work is supported in part by Major State Basic Research Development Program of China (973 Program, 2009CB320903), National Science Foundation (61121002, 61103088), National High-tech R&D Program of China (863 Program, SS2012AA010805) and National Sci-Tech Support Plan (2011BAH08B01).

5. REFERENCES

- [1] M. Ghanbari, *Standard Codes: Image Compression to Advanced Video Coding*. London, U.K.: IEE, 2003.
- [2] P. P. Gandhi, "JPEG-based image compression for low-bit-rate coding," in *Proc. SPIE Still-Image Compression II Conf.*, R. L. Stevenson, A. I. Drukarev, and T. R. Gardos, Eds., vol. 2669, pp. 82–94, 1996.
- [3] M. Yuen and H. R. Wu, "A survey of MC/DPCM/DCT video coding distortions," *Signal Process.*, vol. 70, no. 3, pp. 247–278, Nov. 1998.
- [4] E. Cands, "Compressive sampling," *Proc. Int. Congr. Mathematics*, pp.1433 -1452, 2006.
- [5] Y. Tsaig, M. Elad, G. Golub, and P. Milanfar, "Optimal framework for low bit-rate block coders," in *Proc. IEEE Int. Conf. Image Processing*, vol. II, pp. 219–222, 2003.
- [6] B. Zeng and A. N. Venetsanopoulos, "A JPEG-based interpolative image coding scheme," in *Proc. IEEE ICASSP*, vol. V, pp. 393–396, 1993.
- [7] A. M. Bruckstein, M. Elad, and R. Kimmel, "Down-scaling for better transform compression," *IEEE Trans. Image Process.*, vol. 12, no. 9, pp. 1132–1144, Sep. 2003.
- [8] W. Lin and D. Li, "Adaptive downsampling to improve image compression at low bit rates," *IEEE Trans. Image Process.*, vol. 15, no. 9, pp.2513 -2521, 2006.
- [9] X. Wu, X. Zhang and X. Wang, "Low bit-rate image compression via adaptive down-sampling and constrained least squares upconversion," *IEEE Trans. Image Process.*, vol. 18, no. 3, pp.552 -561, 2009.
- [10] Jian Zhang, Debin Zhao, Chen Zhao, Ruiqin Xiong, Siwei Ma, and Wen Gao, "Image Compressive Sensing Recovery via Collaborative Sparsity", *IEEE Journal on Emerging and Selected Topics in Circuits and Systems, Special Issue on Circuits, Systems and Algorithms for Compressive Sensing*, vol. 2, no. 3, pp. 380–391, Sep. 2012.
- [11] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Trans. on Image Process.*, vol. 18, pp. 2419–2434, 2009.
- [12] J. F. Cai, S. Osher, and Z. W. Shen, "Split Bregman methods and frame based image restoration," *Multiscale Model. Simul.*, vol. 8, pp. 5057-5071, 2009.
- [13] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE. Trans. Image Process.*, vol. 13, no. 4, Apr. 2004.
- [14] J. Zhang, D. Zhao, C. Zhao, R. Xiong, S. Ma, and W. Gao, "Compressed Sensing Recovery via Collaborative Sparsity", Prof. of IEEE Data Compression Conference, pp. 287–296, Snowbird, Utah, USA, Apr. 2012.
- [15] J. Zhang, R. Xiong, C. Zhao, S. Ma, D. Zhao. "Exploiting Image Local and Nonlocal Consistency for Mixed Gaussian-Impulse Noise Removal", Prof. of IEEE Int. Conf. on Multimedia & Expo, pp. 592–597, Melbourne, Australia, Jul. 2012.