A PIPELINE FOR LENSLET LIGHT FIELD QUALITY ENHANCEMENT

Pierre Matysiak, Mairéad Grogan, Mikaël Le Pendu, Martin Alain, Aljosa Smolic

V-SENSE Project, School of Computer Science and Statistics, Trinity College, Dublin

ABSTRACT

In recent years, light fields have become a major research topic and their applications span across the entire spectrum of classical image processing. Among the different methods used to capture a light field are the lenslet cameras, such as those developed by Lytro. While these cameras give a lot of freedom to the user, they also create light field views that suffer from a number of artefacts. As a result, it is common to ignore a significant subset of these views when doing high-level light field processing. We propose a pipeline to process light field views, first with an enhanced processing of RAW images to extract sub-aperture images, then a colour correction process using a recent colour transfer algorithm, and finally a denoising process using a state of the art light field denoising approach. We show that our method improves the light field quality on many levels, by reducing ghosting artefacts and noise, as well as retrieving more accurate and homogeneous colours across the sub-aperture images.

Index Terms— Light Fields, Lenslet Decoding, Colour Transfer, Denoising, Plenoptic RAW Processing

1. INTRODUCTION

Light fields aim to capture all light rays passing through a given amount of the 3D space [1]. Compared to traditional images representing a projection of light rays on a 2D plane, a 4D light field also contains the angular direction of the rays. The light field of a real scene can be captured with different devices such as a single camera on a moving gantry, an array of cameras, or a plenoptic camera including an array of micro-lenses in front of its sensor. The latter has received a lot of attention since the commercialisation by the Lytro company of two successive models capable of capturing light fields with a dense angular sampling. A concurrent plenoptic camera design called plenoptic 2.0 has also been proposed in [2]. Unlike the former design (referred to as unfocused), each micro-lens produces a focused micro-image.

Because of the micro-lens array, the generation of exploitable images from the RAW sensor data is significantly more complex than with traditional cameras. Furthermore, there is no consensus on the light field representation to adopt. While plenoptic 2.0 cameras are generally used to directly render images at varying focus, the unfocused plenoptic cameras are better suited for

the extraction of sub-aperture images (SAI) with a very wide depth of field, each corresponding to a viewpoint of the scene. Since the latter representation is more commonly used for various light field applications (e.g. depth estimation, compression), we consider in our analysis the extraction of SAIs from unfocused plenoptic camera RAW data. Among the different methods proposed for this task [3–6], the most complete pipeline was developed by Dansereau et al. [3]. It has been widely adopted by the light field research community. For example, it has a central role in the standardisation effort for light field compression as it is used as part of the JPEG PLENO [7] test set. However, the extracted views may suffer from many artefacts including noise, unnatural horizontal stripes, ghosting effects on external SAIs, colour and brightness inconsistencies between SAIs, inaccurate colour balance, and important loss of dynamic range. These limitations have a negative impact on most light field applications such as depth estimation, segmentation, rendering or compression (see [8] for a comprehensive overview). Furthermore, external SAIs containing essential depth information are often ignored because of extreme distortions. Note that, although the proprietary Lytro Desktop software overcomes many of these issues, it essentially targets the rendering of refocused images, and it is not suitable for generating an SAI array.

Later research on the subject in [5,6,9–12] has essentially focused on adapting the demosaicing step which retrieves the RGB colour components of each pixel from the partial colour information actually captured by camera sensors. We believe that a more global analysis of the pipeline is also necessary.

In this paper, we present an enhanced processing pipeline for lenslet-based plenoptic cameras. We first propose improvements within the RAW processing of [3]. In particular, we show how the devignetting step (i.e. correction of lenslet vignetting) negatively impacts the overall image aspect (colour balance, brightness, loss of dynamic range), and how to correct it. We additionally recommend the use of white image guided interpolations [9] to reduce the ghosting effect of external SAIs. Based on the observation that brightness and colour inconsistencies between SAIs can hardly be corrected in the early stages of the RAW processing without introducing other artefacts, we perform a post processing colour correction step. We finally analyse the noise level at each step of the process and suggest that additional denoising should be applied preferably after the colour correction.

2. PROPOSED PIPELINE

2.1. RAW Light Field Decoding

In this section, we analyse the early steps of the RAW processing : the lenslet devignetting, the demosaicing, and the interpolations required for compensating misalignments between the micro-lens array and the sensor grid.
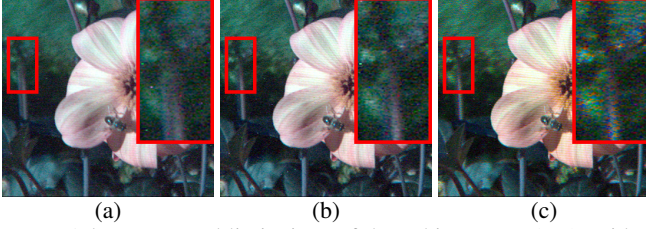
Fig. 1: Advantages and limitations of the White Image (WI) guided method of [9]: (a) standard demosaicing [13] and bicubic interpolations, (b) standard demosaicing [13] and WI-guided interpolations, (c) WI-guided demosaicing and interpolations.

2.1.1. Lenslet Devignetting

In [3], lenset devignetting is performed first as it results in more uniform brightness over the sensor array and thus, easier demosaicing. This step simply consists of a pixel-wise division of the RAW image by a RAW White Image (WI) that exhibits the pattern of micro-lens vignetting. Since the WI was previously taken by the same device as the picture to process, this division step does not only remove the vignetting pattern, but also implicitly normalises the colour responses of the red, green, and blue pixels on the sensor. However, the normalisation is not taken into account by traditional RAW processing [3]. In particular, the white balance parameters, determined by the camera during the capture (either automatically or with user interaction) and stored as metadata, do not apply to RAW data with normalised RGB responses. In order to obtain the intended colour balance, we multiply the red and blue pixels of the WI by normalisation factors provided as metadata of the camera. Note that these factors may also be obtained by colour calibration of the sensor.

Since the pixel values of the WI are lower than 1 even at micro-lens centres, the devignetting in [3] also increases the overall brightness of the light field. Bright areas reaching higher values than 1 after devignetting are considered as saturated in the rest of the process, and the information is lost. Therefore, we also apply a global normalisation of the WI by dividing all the pixels by its 99.9th percentile (we do not use the maximum value to exclude hot pixels).

2.1.2. Demosaicing and Interpolations

Previous analysis by David et al. [9] have shown how standard demosaicing and interpolations introduced both ghosting artefacts and fading of the colours for the external SAIs. In order to reduce the problem, they adapted those steps by weighing the contribution of each pixel using the vignetting pattern of the White Image. Two observations can be made from their results. Firstly, the ghosting effect is essentially reduced by the adaptation of the interpolation step (see Fig. 1(b)). Secondly, while the modified demosaicing improves the overall colour consistency between SAIs, it may also create colour noise (see Fig. 1(c)). Hence, we suggest that only the WI-guided interpolations should be used, and we propose in the next section a post-processing step to enforce colour homogeneity in the light field.

2.2. Colour Correction

To obtain homogeneous colours in the light field, we successively process each SAI by performing colour transfer from a reference SAI designated as palette image. Several propagation schemes defining the palette image for each SAI and the processing order are proposed in section 2.2.2.

For a given pair of target and palette SAIs, our colour correction addresses the viewpoint disparity by incorporating correspondence estimations in the recent colour transfer algorithm of Grogan et al. [14–16] described in section 2.2.1. Their global approach was found to be more robust to erroneous correspondences and outperforms existing colour correction methods [17–19]. In our implementation, coarse-to-fine patch matching (CPM) [20] was chosen to estimate pixel correspondences between the views since it is both accurate and efficient, and has been successfully used as an initialisation step for optical flow computation for light fields by Chen et al. [21]. Once the pixel correspondences are estimated, we pass their colour values to the colour transfer algorithm.

2.2.1. Colour Transfer

Given a set of n colour correspondences $(c_t^{(k)}, c_p^{(k)})_{k=1\ldots n}$ between the target and palette images, where the set of colours $c_t^{(k)}$ from the target image should correspond to the colours $c_p^{(k)}$ from the palette after recolouring, Grogan et al. [15, 16] propose to fit a Gaussian Mixture Model to each set of correspondences as follows:

$$p_t(x|\theta) = \sum_{k=1}^{n} \frac{1}{n} \mathcal{N}(x; \phi(c_t^{(k)}, \theta), h^2 \mathrm{I}) \qquad (1)$$

and

$$p_p(x) = \sum_{k=1}^{n} \frac{1}{n} \mathcal{N}(x; c_p^{(k)}, h^2 \mathrm{I}) \qquad (2)$$

The vector $x \in \mathbb{R}^3$ takes values from a 3D colour space, and each Gaussian is associated with an identical isotropic covariance matrix, $h^2 \mathrm{I}$. The colours $\phi(c_t^{(k)}, \theta)$ are obtained by transforming $c_t^{(k)}$ by some transformation ϕ which depends on θ. The goal is then to estimate the transformation ϕ that registers $p_t(x|\theta)$ to $p_p(x)$, and thus transforms the colour distribution of the target image to match that of the palette image. Grogan et al. propose letting ϕ be a global parametric thin plate spline transformation, and estimate the parameter θ controlling ϕ by minimising :

$$\mathcal{C}(\theta) = -\langle p_t | p_p \rangle = \sum_{k=1}^{n} \frac{1}{n^2} \mathcal{N}(0; \phi(c_t^{(k)}, \theta) - c_p^{(k)}, 2h^2 \mathrm{I}) \quad (3)$$

For our application, we found that using the LAB space representation of colours gave better results than RGB. We also altered the simulated annealing parameters proposed in [15] to ensure that local minima were avoided during optimisation.

2.2.2. Propagation

To guarantee colour homogeneity over the whole light field, we investigate three propagation schemes. The first is a more naive approach in which we recolour every SAI in the light field taking the centre view as palette, as it has the most accurate colours. This involves computing correspondences between each SAI and the centre one and passing them to Eqs. (1) and (2) to perform colour correction. While these correspondences are accurate in many cases, as the disparity increases over the light field, fewer accurate correspondences are available which can affect the quality of the recoloured external views.

To combat this, we investigated a second scheme which involves propagating the colours incrementally starting from the

(a) Original RAW decoding [3]   (b) Proposed RAW decoding   (c) Recoloured

Fig. 2: Overview of all sub-aperture images in a light field. Showcases the differences in appearance between the original and proposed RAW decoding, and the homogeneity of the colours across views after colour correction.

centre view, along the centre column, and then along each row. This is achieved by first taking the centre view as palette and using it to recolour its two outer neighbouring views in the column. Once corrected, these two views are used to recolour their outer neighbouring views. This process is repeated until all views in the column are corrected. Following a similar process, the colours from the centre column are propagated out to each of the rows, with the centre view in each row initialised as palette. This scheme ensures that neighbouring views have very few colour differences between them.

The final scheme is a combination of the previous two, with each view (apart from the centre one) recoloured using the colour correspondences from its already corrected inner column or row neighbour as well as those from the centre view, and combining them when passing them to Eqs. (1) and (2), in order to maintain colour consistency across the light field. The three methods will henceforth be referred to as 'centre', 'prop' and 'prop+centre'.

2.3. Denoising

In addition to the colour artefacts corrected by the previous step of the proposed pipeline, light fields captured with the Lytro Illum are known to exhibit camera noise pattern. We thus propose to apply denoising as a final step. For that purpose, we use the state-of-the art LFBM5D filter introduced in [22]. This filter takes full advantage of the 4D nature of light fields by creating disparity compensated 4D patches which are then stacked together with similar 4D patches along a 5th dimension. These 5D patches are then filtered in the 5D transform domain, obtained by cascading a 2D spatial transform, a 2D angular transform, and a 1D transform applied along the similarities.

Denoising is therefore applied after colour correction, so that the dark corner SAIs can be denoised together with the rest of the light field.

3. RESULTS

We present results of our pipeline applied on a subset of the freely available EPFL [23] and INRIA [24] datasets captured with Lytro Illum cameras. Detailed dataset composition and additional results are also available online[1].

3.1. Colour Quality

We first show in Fig. 3 the importance of the simple normalisation steps proposed in section 2.1.1 for the colour balance and overall brightness. For reference, the bottom right part of each

---

[1]https://v-sense.scss.tcd.ie/?p=1548

sub-figure shows a refocused image obtained by the Lytro proprietary software with the intended colours (i.e. as displayed by the camera when taking the picture). Note that the results of [3] are often wrongly assumed to be gamma corrected, leading to exaggerated contrasts and colour saturation. For a fair comparison, we performed standard sRGB gamma correction for both methods.

Regarding colour correction, we determined the best recolouring scheme (see Sec. 2.2.2) by comparing their results using a number of metrics : S-CIELab [25], a global histogram distance, and a blind noise level estimation [26]. We investigated other metrics for comparison (PSNR, SSIM [27], CID [28]), but found that the three presented here were the most comprehensive. To determine the accuracy of the colour correction with respect to the centre view, we computed the difference between the centre SAI and each SAI in the light field using S-CIELab and the histogram distance, and averaged the results over SAIs. As S-CIELab compares local colour differences between images, the disparity between SAIs may affect the evaluation. However, as all methods are compared on the same set of light fields with the same disparity differences, values are still indicative of colour correction accuracy. The global histogram metric on the other hand is more robust to these disparity changes. For a pair of images, the histogram distance is measured as the average χ-square differences between their L, a*, and b* histograms, each computed on 25 bins.

From Fig. 6 we can see that all approaches improve the overall colour similarity between the centre view and the remainder of the light field, with the 'centre' scheme marginally capturing the colours of the centre view more faithfully according to S-CIELab and the histogram distance. However, as previously mentioned, it displays inaccuracies in some unfavourable cases, when the scene contains plain textureless colours and accurate correspondences are difficult to compute (e.g. Color_Chart in Fig. 4). On the other hand, the 'prop' and 'prop+centre' schemes

(a) Method [3] vs Lytro Desktop   (b) Our method vs Lytro Desktop

Fig. 3: Above red line: central SAI of the light field obtained with (a) Dansereau et al. [3], (b) Our method. Below red line: refocused image from Lytro Desktop proprietary software (using 'as shot' white balance option).

Fig. 4: Recolouring examples. First column shows the centre SAI (red and blue lines are used to create the EPIs in Fig. 5); second column is an SAI picked on the border of the light field; third and fourth columns are the same after recolouring using the 'centre' and 'prop+centre' schemes respectively.

(a)

(b)

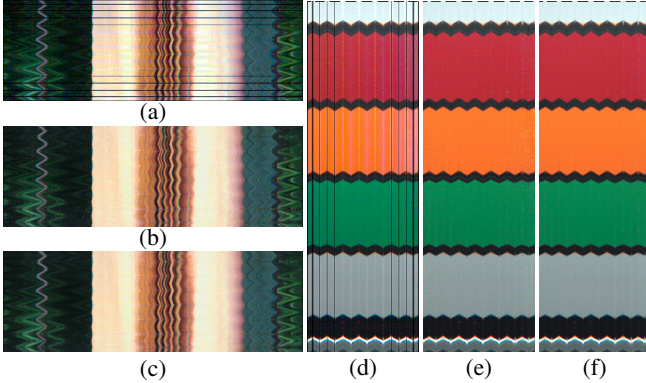(c)   (d)   (e)   (f)

Fig. 5: Stacked epipolar images showcasing colour differences in the LFs Bee_2 (a,b,c) and Color_Chart (d,e,f): after our RAW decoding (a,d), after 'prop+centre' recolouring (b,e), and after denoising (c,f). Dark lines in (a,d) are caused by the dark SAIs in the corner of the light field (we only excluded the most extreme ones which are completely black and cannot be corrected). Selected lines are shown in Fig. 4. **As with Fig. 4, best viewed in colour and zoomed in.**

produce comparable results to 'centre' with respect to S-CIELab and the histogram distance, and create less noise (Fig. 7). They also ensure that neighbouring SAIs have consistent colours. While we found that using the 'prop+centre' scheme can sometimes give a more accurate colour correction compared to 'prop', using additional correspondences increases the computational complexity of the colour correction step, and the decision to use these additional correspondences is left to the user.

We visually assess the results of our recolouring method in Figs. 2, 4 and 5. The results are visually pleasing, with smooth transitions between consecutive views, seen in Fig. 2, and the colours overall remaining consistent with those in the centre view (see also Fig. 4). This is particularly visible when computing epipolar images (as seen in Fig. 5), which consist of stacks of the same horizontal or vertical line of pixels taken across all the views of the light field. These images show a clear improvement in keeping the colours consistent over the whole light field, which is further improved after the denoising process. Fig. 2 shows that our colour correction also successfully recolours the dark corner images in the light field, which can then be taken advantage of by other processing tools.

3.2. Noise Analysis

In order to quantify the noise reduction, we perform blind noise level estimation [26] before and after denoising. Assum-
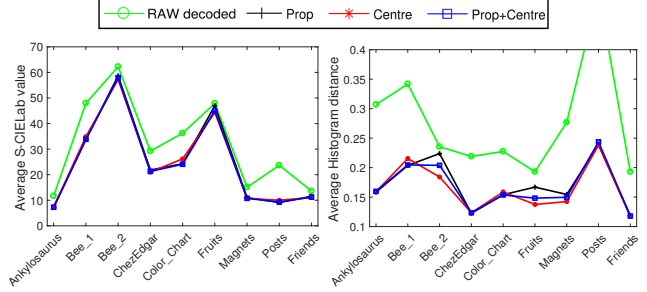

Fig. 6: Metric comparison, using S-CIELab [25] and histogram distance. Lower values are better. It shows all three schemes are outputting comparable results.
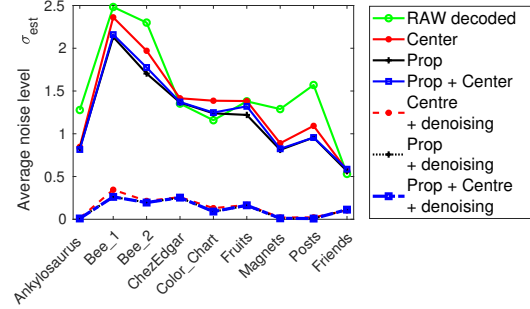

Fig. 7: Noise level estimated for each light field using [26], before and after denoising.

ing an Additive White Gaussian Noise (AWGN), we estimate the noise standard deviation σ_{est} for each SAI, and then compute the average over the whole light field. Although the AWGN model is a simplification of the actual noise, our intent here is to provide a relative comparison of the different schemes rather than an absolute noise measure. We report in Fig. 7 the estimated values for each light field. Results show that overall the colour correction step can slightly decrease the noise level. However, we observe that in some cases the noise is amplified (e.g. Color_Chart), which further justifies applying denoising last. The noise level is clearly reduced for all approaches after applying the denoising. As mentioned previously, the 'centre' scheme exhibits a higher noise level than the other two tested approaches, even after denoising. A visual comparison before and after denoising is shown in Fig. 5.

4. CONCLUSION

We have presented a pipeline which aims at substantially improving upon the overall visual quality of SAIs in a light field. The final results show that every processing step provides necessary and complementary benefits. We feel that providing a way to enhance the available lenslet camera datasets is necessary as such a complete approach does not currently exist. Visual inspection as well as metric comparison show that our method provides significant improvement on the quality of the light field views, counteracting the unfortunate side-effects lenslet cameras suffer from. Furthermore, by allowing the use of a higher number of views when performing high-level light field processing, such as depth estimation, segmentation or rendering, we hope to improve the results of those algorithms.

5. REFERENCES

[1] M. Levoy and P. Hanrahan, "Light field rendering", in *Proc. SIGGRAPH*, 1996, pp. 31–42.

[2] A. Lumsdaine and T. Georgiev, "The focused plenoptic camera", in *Proc. ICCP*, 2009, pp. 1–8.

[3] D. G. Dansereau, O. Pizarro, and S. B. Williams, "Decoding, calibration and rectification for lenselet-based plenoptic cameras", in *Proc. CVPR*, 2013, pp. 1027–1034.

[4] D. Cho, M. Lee, S. Kim, and Y. W. Tai, "Modeling the calibration pipeline of the lytro camera for high quality light-field image reconstruction", in *Proc. ICCV*, Dec. 2013, pp. 3280–3287.

[5] S. Xu, Z.-L. Zhou, and N. Devaney, "Multi-view image restoration from plenoptic raw images", in *Proc. ACCV Workshops*, 2014.

[6] M. Seifi, N. Sabater, V. Drazic, and P. Perez, "Disparity-guided demosaicking of light field images", in *Proc. IEEE ICIP*, Oct. 2014, pp. 5482–5486.

[7] "JPEG Pleno call for proposals on light field coding", in *ISO/IEC JTC 1/SC29/WG1N73013*, Chengdu, China, Oct. 2016.

[8] G. Wu, B. Masia, A. Jarabo, Y. Zhang, L. Wang, Q. Dai, T. Chai, and Y. Liu, "Light field image processing: An overview", *IEEE Journal of Selected Topics in Signal Processing*, vol. 11, no. 7, pp. 926–954, Oct. 2017.

[9] P. David, M. Le Pendu, and C. Guillemot, "White lenslet image guided demosaicing for plenoptic cameras", in *Proc. IEEE MMSP*, Oct. 2017, pp. 1–6.

[10] T. Lian and K. Chiang, "Demosaicing and denoising on simulated light field images", 2016.

[11] Z. Yu, J. Yu, A. Lumsdaine, and T. Georgiev, "An analysis of color demosaicing in plenoptic cameras", in *Proc. CVPR*, Jun. 2012, pp. 901–908.

[12] X. Huang and O. Cossairt, "Dictionary learning based color demosaicing for plenoptic cameras", in *Proc. CVPR Workshops*, Jun. 2014, pp. 455–460.

[13] H. S. Malvar, L.-W. He, and R. Cutler, "High-quality linear interpolation for demosaicing of bayer-patterned color images", in *Proc. IEEE ICASSP*, 2004.

[14] M. Grogan, M. Prasad, and R. Dahyot, "L2 registration for colour transfer", in *Proc. EUSIPCO*, Aug 2015, pp. 1–5.

[15] M. Grogan and R. Dahyot, "Robust registration of gaussian mixtures for colour transfer", ArXiv e-prints (May 2017). arXiv:cs.CV/1705.06091.

[16] M. Grogan, R. Dahyot, and A. Smolic, "User interaction for image recolouring using L2", in *Proc. CVMP*, 2017, pp. 6:1–6:10.

[17] Y. Hwang, J. Y. Lee, I. S. Kweon, and S. J. Kim, "Color transfer using probabilistic moving least squares", in *Proc. CVPR*, Jun. 2014, pp. 3342–3349.

[18] N. Bonneel, J. Rabin, G. Peyré, and H. Pfister, "Sliced and radon wasserstein barycenters of measures", *Journal of Mathematical Imaging and Vision*, vol. 51, no. 1, pp. 22–45, Jan 2015.

[19] F. Pitié, A. C. Kokaram, and R. Dahyot, "Automated colour grading using colour distribution transfer", *Computer Vision and Image Understanding*, vol. 107, no. 1, pp. 123 – 137, 2007, Special issue on color image processing.

[20] Y. Hu, R. Song, and Y. Li, "Efficient coarse-to-fine patchmatch for large displacement optical flow", in *Proc. CVPR*, 2016, pp. 5704–5712.

[21] Y. Chen, M. Alain, and A. Smolic, "Fast and accurate optical flow based depth map estimation from light fields", in *Irish Machine Vision and Image Processing Conference (IMVIP)*, 2017.

[22] M. Alain and A. Smolic, "Light field denoising by sparse 5D transform domain collaborative filtering", in *Proc. IEEE MMSP*, Oct. 2017, pp. 1–6.

[23] M. Rerabek and T. Ebrahimi, "New Light Field Image Dataset", in *Proc. QoMEX*, 2016.

[24] "Inria lytro illum dataset", `http://www.irisa.fr/temics/demos/lightField/CLIM/DataSoftware.html`, accessed: 26-01-2018.

[25] X. Zhang and B. A. Wandell, "A spatial extension of CIELAB for digital color-image reproduction", *Journal of the Society for Information Display*, vol. 5, no. 1, pp. 61–63, 1997.

[26] X. Liu, M. Tanaka, and M. Okutomi, "Single-image noise level estimation for blind denoising", *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 5226–5237, Dec. 2013.

[27] 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, pp. 600–612, Apr. 2004.

[28] I. Lissner, J. Preiss, P. Urban, M. S. Lichtenauer, and P. Zolliker, "Image-difference prediction: From grayscale to color", *IEEE Trans. Image Process.*, vol. 22, no. 2, pp. 435–446, Feb. 2013.