Perceptual Comparison of Demosaicing Algorithms and In-camera Demosaicing with JPEG Compression

Bartolomeo Montrucchio
Dipartimento di Automatica e Informatica, Politecnico di Torino, c.so Duca degli Abruzzi 24, Torino, Italy

Keywords: Demosaicing Algorithms, Perceptual Comparison, Color Filter Array (CFA).

Abstract: Color image acquisition in digital cameras is often performed by using CCD or CMOS sensor chips with a color filter array on the top of a single monochromatic sensor. In this paper, a perceptual comparison is performed among three well known demosaicing algorithms plus in-camera demosaicing with lossy compression JPEG, by means of subjective tests, that is with the help of human beings. The novelty of the approach is that chosen algorithms have been selected as representative of those used in commercial raw image converters used by professionals in graphics and that the test has been performed on a large number of people, achieving results only partially similar to the results got by means of computed metrics. The results show that in the greatest part of conditions and for non particularly expert users, the capability of the most advanced demosaicing algorithms of producing an almost perfect reconstruction on the full-color image is not strictly required. Only for selected categories of images it is possible to find a clear winner among the algorithms.

1 INTRODUCTION

Full color images in digital cameras are often obtained putting a Color Filter Array (CFA) in front of a single monochromatic sensor. There exist at least one solution able to avoid CFA, the Foveon X3 chip, that works in a way similar to classical color films, but the greatest majority of digital cameras still uses CFA based sensors.

Reconstructed image can be subjected to several artifacts, and different demosaicing algorithms try to address this problem.

The basic idea of this paper is (as also suggested in (Gunturk et al., 2005)) to verify if from the perceptual point of view is there a true difference among different demosaicing algorithms, characterized by different cpu time for the demosaicing and by different levels of objective and subjective performance. In the paper three different demosaicing algorithms will be subjectively compared among themselves and with the in-camera demosaiced and lossy compressed JPEG produced by a semiprofessional high end Full Frame Canon reflex 5D MarkII camera. The novelty of the approach is based both on the use of an easily available and highly efficient implementation of such demosaicing algorithms and on the wide number of the observers participating to the experiment with respect to similar previous studies. Time and effort required to the final user to process and archive the images with respect to the final subjective quality (that is more interesting to the photographer) have in fact to be considered in the design of the demosaicing algorithm itself.

The rest of the paper is organized as follows. In Section 2 a brief background on demosaicing algorithms is presented. In Sections 3 and 4 the proposed experiment is presented and the results are summarized from a perceptual point of view. Finally in Section 5 conclusions are presented.

2 BACKGROUND

Modern high end digital cameras are always able to store in the memory card not only an already demosaiced image (usually with lossy compression (JPEG) or Tiff format) but also the raw image directly taken from the sensor. Raw images reflect the used CFA, which is usually structured on Bayer's idea. In particular in (Hao et al., 2011) a new method for designing optimal color filter arrays is presented; results are very interesting, since it is possible to design specific CFAs. Many demosaicing algorithms have been developed in the previous years (Gunturk et al., 2005) and a number of new algorithms have

Montrucchio B..
Perceptual Comparison of Demosaicing Algorithms and In-camera Demosaicing with JPEG Compression.
DOI: 10.5220/0004297801300133
In *Proceedings of the International Conference on Computer Vision Theory and Applications* (VISAPP-2013), pages 130-133
ISBN: 978-989-8565-47-1

been presented recently. A basic algorithm can be considered as based on bilinear interpolation of near values for getting missing values; it can be seen e.g. in (Sakamoto et al., 1998). Another common algorithm is (Chang et al., 1999), in which a threshold-based variable number of gradients is used. Gradients are computed in eight directions instead then on a single direction. AHD algorithm (Hirakawa and Parks, 2005) does the demosaicing trying to interpolate in the direction with fewer color artifacts and addressing aliasing with filterbank techniques; moreover interpolation artifacts are reduced by means of nonlinear iterative procedure. An extensive survey can be found in (Li et al., 2008); in this survey directions for future research are outlined. Generally speaking, in order to compare demosaicing algorithms objective and subjective measures can be used. As reported in (Gunturk et al., 2005) objective measures like mean square error (MSE), CIELAB and its spatial extensions (e.g S-CIELAB), and the measure of the zipper effect as "an increase in color difference with respect to its most similar neighbor") are very accurate. Subjective measures are less frequently used. In (Rajeev Ramanath et al., 2002) subjective tests are used to assess numerical measures. However, as reported in (Longere et al., 2002), the use of objective metrics, even if very accurate, can provide results different from the perceptual ones. For these reasons in this paper a fully perceptual approach is proposed.

3 PROPOSED EXPERIMENT

The basic idea of the experiment is to compare Bilinear(Sakamoto et al., 1998), VNG(Chang et al., 1999), AHD(Hirakawa and Parks, 2005) and the in-camera demosaiced JPEG (Canon reflex cameras' JPEG), from a perceptual point of view since, as explained in Section 2, a perceptual comparison can be useful for a better (and global) understanding of the behavior of a demosaicing algorithm. These four different algorithms have been chosen in order to select a basic demosaicing algorithm (Bilinear), a medium strength algorithm (VNG) and a high quality algorithm (AHD), in comparison to the in-camera demosaiced and lossy compressed JPEG provided directly by the camera (usually used when the highest quality is not strictly required). In-camera demosaiced and lossy compressed JPEG (at maximum quality) was included in order to try to understand if developing a raw image with advanced demosaicing algorithms like those included in the software used by professionals for developing raw images (Dcraw (Coffin, 2011) was chosen because its code is probably used as starting point for several very common softwares and because its algorithms, even not the best of all, are anyway very good) is really useful or not.

The images have been acquired by means of a semiprofessional digital camera Canon 5D MarkII (21 Megapixels, full frame 24x36), equipped with professional high quality (even if non diffraction limited) Canon lenses.

Four test images (scenes) have been acquired. In particular in this experiment were involved much more people than in (Longere et al., 2002) (47 instead of 10) and the test was much shorter (in the other case one hour was required). Even with a short test, many volunteers were not coherent when the test was repeated changing the order of the algorithms in the images (procedure not followed in (Longere et al., 2002), so it is not possible to know how much their volunteers were coherent).

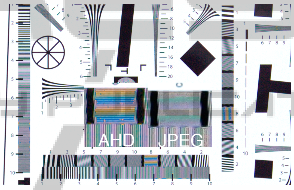

Figure 1: ISO 12233 chart image. Enlarged details are for AHD and in-camera JPEG.

Figure 2: Macro image. Enlarged details are for AHD and in-camera JPEG.

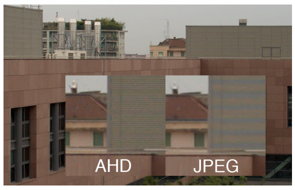

Figure 3: Urban image. Enlarged details are for AHD and in-camera JPEG.

The four images are reported in Fig.1, Fig.2, Fig.3, and Fig.4 and represent a ISO 12233 test chart (technical image, Canon EF 24-105 f4 L IS USM, 50mm, f4, 1/30s, tripod, manual focus with LiveView), a macro-photo (Canon MP-E 65mm f2.8 1-5x used at 3x, f5.6, 1/200s, flash, manual focus with LiveView, the field in the cropped photo is about 3mmx2mm), a typical urban image (Canon 24-105

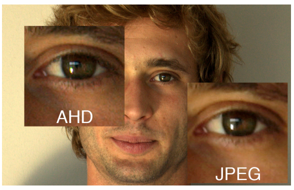

Figure 4: Human face image. Enlarged details are for AHD and in-camera JPEG.

as before used at 24 mm, tripod, 1/200s, f8, manual focus with LiveView) and a human face (Canon 24-105 as before used at 55mm, 1/50s f5.6, autofocus). The experimental setup is based on a high quality 1900x1200 pixel monitor (17" MacBookPro notebook computer) and the four algorithms (each image being of 960x600 pixels) are presented together (given a specific photo) in the same image in order to give to the subject a direct way to compare results. As in (Longere et al., 2002) the perceptual question was which was the most pleasant and detailed image. In this case also the least pleasant and detailed one was requested. The time left to the subject for the image was about 30 seconds.

The involved volunteers were 47, with a mean age of 30 years, mainly males. The original raw and jpg images have been cropped (without scaling) and adjusted in terms of levels, contrast and sharpening as described in (Longere et al., 2002) in order to put all algorithms on the same level of sharpness and colors. All original raw images and test images are available as soon as an internal report more extended with several more details[1]. Volunteers were also asked to declare eyesight problems (color issues were not present). Volunteers involved in eyesight problems were about half of the total. Finally volunteers were asked to assess their level in expertise in digital image photography. Each volunteer was involved in only one session, in which he/she was asked to give a preference on eight images, that is two slides for each photo. In the second slide the order of the four algorithms used to process the photo was changed in order to verify if the volunteer is able to give again the same preference. The test requested no more than fifteen minutes for each person. It is important to note that it was possible also to choose no difference.

4 PERCEPTUAL RESULTS

In this Section the perceptual results will be presented, in order to show how data have been collected. In Fig.5 are presented the results for ISO 12233 chart image, in Fig.6 are presented the results for the macro photo, in Fig.7 are presented the results for the urban photo, and finally in Fig.8 are presented the results for the human face photo. People who had eyesight problems (corrected) showed a result very similar to the people without problems. For this reason specific results will be omitted here. From the perceptual results just reported, it appears that the best (with objective metrics) algorithm, AHD, is not so good from a perceptual point of view. In fact, even if seeing accurately Fig.1, Fig.2, Fig.3 and Fig.4, it is possible to say that AHD is normally better than the other (probably with some doubt in ISO 12233 chart image), results of the perceptual comparison say that AHD is for sure better than the other only in the urban image. In fact in the chart image the best algorithm appears to be JPEG, with a good consistency (Fig.5) between slides 1 and 2; the reason is probably that JPEG, being in this case a demosaicing plus a lossy algorithm, smooths high frequencies, reducing aliasing. Please note that volunteers have told that even without colored aliasing they would have chosen again JPEG, generally. In the macro image there is a great confusion, because AHD wins only in slide 2, there is very few coherency between slide 1 and 2 and the people indicating that there is no differences among the images are more than 10 (Fig.6). In the urban image, as told before, AHD wins very clearly. On 47 people, 40 are able to see the superiority of AHD. The reason is probably that the aliasing and the lack of details of the other algorithms make the comparison easy (Fig.7). For the human face image the winner is again AHD (Fig.8), but it is not a clear winner, because only around 50% of the people choose AHD. The most interesting results are with the people that were coherent between slide 1 and slide 2. This double test, simply mixing the order of the images, has been able to find true expert people, or at least people able to find a perceptual meaning in the images. In macro and human face photos only half or less of the volunteers were able to find again the best image, while in the urban image AHD was also easy to be found and in the ISO chart again volunteers did not change their opinion (in this case choosing mainly JPEG).

5 CONCLUSIONS

A perceptual comparison of three demosaicing algorithms and in-camera integrated demosaicing (with lossy compression) has been performed. The subjective experiments have shown, involving 47 volun-

[1]http://staff.polito.it/bartolomeo.montrucchio/VISAPP2013/

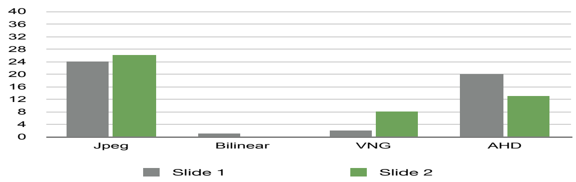

Figure 5: Number of people who chose each algorithm picking the most detailed image given ISO 12233 chart image (slide 1 and slide 2).

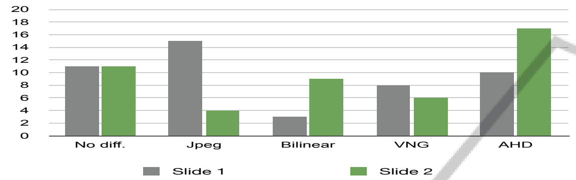

Figure 6: Number of people who chose each algorithm picking the most detailed image given macro image (slide 1 and slide 2).

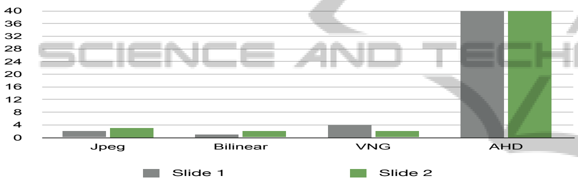

Figure 7: Number of people who chose each algorithm picking the most detailed image given urban image (slide 1 and slide 2).

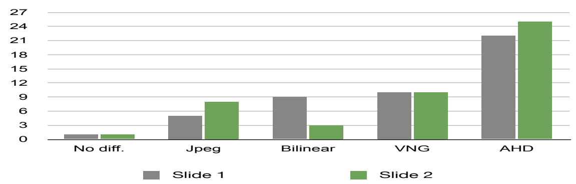

Figure 8: Number of people who chose each algorithm picking the most detailed image given human face image (slide 1 and slide 2).

teers, that the most powerful demosaicing algorithm here considered, AHD, is seen as the clear winner only in very selected cases. In-camera demosaiced (and compressed) JPEG often offers good results, of course uses less space on disk, does not require additional cpu time and for some kinds of images appears even better to general digital image users (probably for the reflex internal processing).

The main result is then that for the great majority of applications, mainly consumer (in this paper diffraction-limited images like those from telescope or microscope have not been considered), in-camera demosaiced and compressed (high quality) JPEG is a very good choice in order to reduce time and effort required to produce and archive the final image. This does not mean that a in-camera demosaiced with lossy compression file contains more details or it is better in an objective way; it means that for several applications (usually consumer) it appears good (and sometimes even better) in a subjective way, allowing to reduce time and effort in developing raw files.

ACKNOWLEDGEMENTS

Many thanks are due to Luca Moglia for conducting the experiments in his M.S. thesis work. Many thanks also to everyone that participated in the experiments and to Bruno Monastero for his important help in statistical analysis.

REFERENCES

Chang, E., Cheung, S., and Pan, D. Y. (1999). Color filter array recovery using a threshold-based variable number of gradients. In *Proc. SPIE*, volume 3650, pages 36–43.

Coffin, D. J. (2011). Dcraw, decoding raw digital photos in linux. http://www.cybercom.net/~dcoffin/dcraw/.

Gunturk, B., Glotzbach, J., Altunbasak, Y., Schafer, R., and Mersereau, R. (2005). Demosaicking: color filter array interpolation. *Signal Processing Magazine, IEEE*, 22(1):44 – 54.

Hao, P., Li, Y., Lin, Z., and Dubois, E. (2011). A geometric method for optimal design of color filter arrays. *Image Processing, IEEE Transactions on*, 20(3):709 –722.

Hirakawa, K. and Parks, T. (2005). Adaptive homogeneity-directed demosaicing algorithm. *Image Processing, IEEE Transactions on*, 14(3):360 –369.

Li, X., Gunturk, B., and Zhang, L. (2008). Image demosaicing: a systematic survey. In *Proc. SPIE*, volume 6822.

Longere, P., Zhang, X., Delahunt, P., and Brainard, D. (2002). Perceptual assessment of demosaicing algorithm performance. *Proceedings of the IEEE*, 90(1):123 –132.

Rajeev Ramanath, A., Snyder, B., and Hinks, C. (2002). Image comparison measure for digital still color cameras. In *Image Processing. 2002. Proceedings. 2002 International Conference on*, volume 1, pages I–629 – I–632 vol.1.

Sakamoto, T., Nakanishi, C., and Hase, T. (1998). Software pixel interpolation for digital still cameras suitable for a 32-bit mcu. *Consumer Electronics, IEEE Transactions on*, 44(4):1342–1352.