Reconstruction of High-Resolution Images from Asymmetrically-Sampled Sensor Data

D. Vernon[1], G. Doemens[2], and N. Murphy[2]

[1] Department of Computer Science, National University of Ireland, Maynooth
[2] Siemens AG, Munich, Germany
[3] Department of Electrical Engineering, Dublin City University, Ireland

Abstract. The yield of the CMOS or CCD fabrication processes typically places a upper-bound on the resolution with which sensors can be cost-effectively manufactured. In this paper, we present an innovative approach which directly addresses this difficulty by adopting a non-uniform photosite geometry. In particular, we present a technique to reconstruct a high-resolution image by combining two lower-resolution and images which exhibit asymmetric photosite aspect ratios and asymmetric sampling density. This has been done for two configurations of the sensor geometry, one where the photosite centroids are co-linear in both directions and another where the photosite centroids are co-linear in the direction of greater sampling density (i.e. along the sensor columns) but where the origin of each column is offset (staggered) by half the column inter-photosite distance. The technique invokes assumptions of local smoothness to provide the additional constraints required to compute the high-resolution pixel values. Good qualitative and quantitative results have been achieved with a 23% improvement in the root of mean square (RMS) error over a simple reconstruction-based on local averaging.

Keywords: image sampling; CMOS sensors; sensor geometry; image reconstruction

1 Introduction

The need for increased accuracy in visual inspection and metrology, particularly in the manufacture of printed circuit boards, makes the fabrication of high-resolution 2-D image sensors very desirable. However, significantly-increased resolution presents two principal difficulties.

The first difficulty concerns the quadratic growth in the amount of image data generated as resolution is generated. This in turn places a sometimes unsustainable load on image acquisition, processing, and analysis sub-systems. The normal solution to this problem is to increase both the communication bandwidth between these sub-system and the processing power of the system responsible for the interpretation of the image data [1]. Alternative approaches has also been developed that adopt a strategy of increased efficiency, rather than increased power, relying on the

construction of sensors whose photosites or pixels are randomly addressable and relying on the development of processing and analysis algorithms which can exploit this random access feature to effect efficient image interpretation [1-7]. This alternative strategy is particularly attractive when one considers the strict time and resolution constraints imposed by the use of standard video interfaces, such as CCIR and NTSC. Specifically, these standards impose a lower limit on the time which must elapse between the acquisition of each image (40 ms in the case of CCIR) and an upper limit on the effective resolution of the image (576 lines in the case of CCIR).

The second difficulty concerns the yield of the CMOS or CCD fabrication process and the consequent ability to produce image sensors which are free of flaws (or, more usually, sensors with an acceptable number of isolated flaws). As the resolution, and the number of pixels, of the sensor increases, the effective yield drops significantly. Typically, this has presented a substantive barrier to the production of cost-effective high-resolution sensors.

In this paper, we present an innovative approach which directly addresses the second of the above two difficulties (*i.e.* reduced yield) by adopting a non-uniform photosite geometry. The approach also facilitates very efficient solutions to the first difficulty caused by increased size of the data sets through the use of random access strategies. However, we will concentrate on the latter problem in this paper.

2 The Design the RAMAP Sensor

Our sensor design criteria are to produce a sensor geometry which can provide:

(a) the highest possible sensor resolution for application specific problems
(b) the highest possible fabrication yield
(c) the lowest cost
(d) and, significantly, the highest possible sensor resolution for general purpose vision tasks.

These criteria are set out in order of importance. In this instance, the application specific problems require highly-accurate localization of image features in two orthogonal directions. Consequently, it is necessary for the sensor to have a high sampling density in one direction only, assuming one can deploy two sensors to effect the localization. This means that one can exploit a sensor with asymmetric sampling densities (*i.e.* higher in one direction than another) and consequently reduce the number of photosites, in turn increasing the possible yield and reducing the overall cost. For example, see figure 1. However, such a strategy can pose some difficulties with criterion (d) above which is concerned with general-purpose visual analysis. This criterion requires the highest possible resolution in all directions, necessitating a symmetric and high sampling density. In this paper, we show how one can combine the images generated from two asymmetric sensors to reconstruct an image with high resolution and we do this for two configurations of the sensor geometry, one where the photosite centroids are co-linear in both directions (see figure 1.a) and another

where the photosite centroids are co-linear in the direction of greater sampling density (*i.e.* along the sensor columns) but the origin of each column is offset (staggered) by half the column inter-photosite distance (see figure 1.b).

The resolution of the RAMAP sensor is 600x1800 photosites in both cases.

3 Reconstruction

The goal of the work described in this paper is to combine two images derived from either of the sensor configurations shown in figure 1 with another equivalent sensor rotated by ninety degrees with respect to the other. Since the aspect ratio of the photosites is 3:1 and the distribution of photosites across the sensor surface is in the ratio of 1:3 per unit distance, this means that a sensor of $\frac{n}{3} \times n$ pixels has an effective 1:1 aspect ratio. Consequently, it is possible to superimpose two images ($\frac{n}{3} \times n$ and $n \times \frac{n}{3}$) derived from orthogonally-oriented sensors. Thus, we wish to combine a $\frac{n}{3} \times n$ image and a $n \times \frac{n}{3}$ image to yield a $n \times n$ image which exhibits an effective resolution greater than one would generate from simply superimposing and averaging the two images.

Clearly, there are two cases to consider here: (a) the combination of the co-linearly arranged (unstaggered) photosites sensor images and (b) the combination of the staggered photosite sensor images. We will treat each of these in turn.

4 Combination of Unstaggered Photosite Sensor Images

The direct superposition of the unstaggered photosite sensor images yields sub-regions where a 1×3 subregion in one image corresponds exactly to a 3×1 sub-region in the other (see Fig. 2). We wish to re-construct a 3×3 sub-region from these two sub-regions. Thus, the problem reduces to one of estimating the nine pixel values in the 3×3 sub-region on the basis of six pixels values comprising the 1×3 and 3×1 sub-regions. We label the nine required pixel values $p1$ to $p9$ and the six known pixel values a, b, and c, and d, e, and f, in the 1×3 and 3×1 sub-regions respectively (see Fig 3). Since pixel a represents the average light intensity incident on pixels 1, 4, and 7, and similar relations exist for pixels, b through f, we can immediately construct six equations as follows:

$$a = (p1 + p4 + p7) / 3 \quad (1)$$

$$b = (p2 + p5 + p8) / 3 \quad (2)$$

$$c = (p3 + p6 + p9) / 3 \quad (3)$$

$$d = (p1 + p2 + p3) / 3 \quad (4)$$

$$e = (p4 + p5 + p6) / 3 \quad (5)$$

$$f = (p7 + p8 + p9) / 3 \quad (6)$$

However, these equations are not all linearly independent. We can see this by considering the fact that we can show from the above that the sum of a, b, and c is equal to the sum of d, e, and f and hence it is possible to compute any of a, b, c, d, e, or f, if the other five pixel values are known. Consequently, we have five equation and nine unknowns (pixels $p1$–$p9$), implying that we require another four equations. We create these by invoking smoothness constraints on the image. Specifically, we require that the partial derivatives of the image function $\frac{\partial f(x,y)}{\partial x}$ and $\frac{\partial f(x,y)}{\partial y}$ are locally constant. Thus:

$$f(x+1,y) - f(x,y) = k1 = f(x+2,y) - f(x+1,y)$$

and

$$f(x+1,y+1) - f(x,y+1) = k2 = f(x+2,y+1) - f(x+1,y+1)$$

Using our pixel labels, we have:

$$p2 - p1 = p3 - p2$$
$$p5 - p4 = p6 - p5$$

Subtracting and re-arranging, we have:

$$p5 - p2 = p4 - p1 + p6 - p3 - p5 + p2$$

Hence:

$$2(p5 - p2) = p4 - p1 + p6 - p3$$

Adding $(p5 - p2)$ to both sides and re-arranging:

$$3(p5 - p2) = (p4 + p5 + p6) - (p1 + p2 + p3)$$

Thus:

$$p5 - p2 = (p4 + p5 + p6)/3 - (p1 + p2 + p3)/3$$

$$p5 - p2 = e - d \tag{7}$$

Similarly:

$$p5 - p4 = b - a \tag{8}$$

$$p6 - p5 = c - b \tag{9}$$

The ninth equation (remember one of the above is linearly dependent) comes from the estimate of the partial derivative in the direction of the line joining pixels $p1$ and $p5$ (*i.e.* in the direction given by the angle $\frac{\pi}{4}$. On the one hand, this is given straightforwardly by $p5 - p1$. On the other, it is also given by the component in this direction of the gradient magnitude, *viz:*

$$\sqrt{\left(\frac{\partial f(x,y)}{\partial x}\right)^2 + \left(\frac{\partial f(x,y)}{\partial y}\right)^2}\cos\left(\theta - \frac{\pi}{4}\right)$$

where:

$$\theta = \arctan\left(\frac{\partial f(x,y)}{\partial y} \Big/ \frac{\partial f(x,y)}{\partial x}\right)$$

Hence,

$$p5 - p1 = \sqrt{(p5 - p4)^2 + (p5 - p2)^2}\cos\left(\theta - \frac{\pi}{4}\right) \tag{10}$$

where

$$\theta = \arctan\left(p5 - p2 \Big/ p5 - p4\right)$$

Rearranging equations (1) to (10), we get the following expressions for each of the pixels $p1$ to $p9$.

$$p4 = (2a - b - c + 3e)/3 \tag{11}$$

$$p6 = p4 - a + c \tag{12}$$

$$p5 = p4 + b - a \tag{13}$$

$$p2 = p5 - e + d \tag{14}$$

$$p8 = 3b - p2 - p5 \tag{15}$$

$$p1 = \sqrt{(p5 - p4)^2 + (p5 - p2)^2}\cos\left(\theta - \frac{\pi}{4}\right) \tag{16}$$

$$p3 = 3d - p2 - p1 \tag{17}$$

$$p7 = 3a - p4 - p1 \tag{18}$$

$$p9 = 3c - p6 - p3 \tag{19}$$

Recall that equations (11) to (19), which used to compute the nine pixel values in the recontructed image, are based on the assumption that the partial derivatives in the x and y directions are locally constant (over a 3×3 neighbourhood). For the most part, this is a valid assumption. However, there are some circumstances where we know that it will lead to a poor approximation. Specifically, if the partial derivatives change sign (and remember that we have two estimates of each partial derivative in each region) then certainly the assumption is violated. In this case, it is likely that the estimate of $p1$ given by equation (16) will be poor and using equations (17), (18), and (19) to compute $p3$, $p7$, and $p9$ only serves to propagate the error. In such cases of sign change, we instead estimate $p3$, $p7$, and $p9$ independently using the appropriate partial derivatives in a similar way to that given in equation (16) for $p1$.

5. Combination of Staggered Photosite Sensor Images

Unfortunately, the direct superposition of the staggered photosite sensor images does not yield an exact spatial correspondence between the 1×3 and 3×1 sub-regions (see figure 4). We have overcome this by transforming the staggered image into an unstaggered image by estimating the unstaggered pixel values in every alternate row (or column) of the staggered image. Simple linear interpolation is used to compute the estimate (see figure 5). The reconstruction then proceeds exactly as in the case of the unstaggered images, as set out above. As one would expect, this intermediate re-approximation distorts the image somewhat and consequently, the reconstructed image is less faithful than in the case of the unstaggered sensor geometry. This is borne out by the results which are summarized in figures 15 and 16 and Tables 1 and 2 .

6 Results and Evaluation

In order to assess the technique for staggered and unstaggered geometries described above, the algorithm was tested on simulated sensor data. Figure 6 shows a uniformly sampled grey-level image of the first of five test scenes. Figures 7 and 8 show the simulated images for the unstaggered geometry with the sensor orientation at zero and ninety degrees, respectively. Figure 9 shows the effect of reconstituting the image by simple averaging and figure 10 shows the effect of using the model presented in this paper. Figures 11 through 14 show these results for the same scene but with the staggered geometry.

The quality of the reconstruction is estimated by computing the root mean square (RMS) of the error in pixel grey-levels estimates in each case. These figure of merit values are given for all five scenes (see figures 6 and 15) in Tables 1 & 2 for unstaggered and staggered sensor geometries, respectively. These values are shown graphically in Figures 16 and 17.

It is clear from the results presented here that the RMS error is considerably smaller for those images which do not exhibit a preponderance of intensity discontinuities (*e.g.* scenes 1, 3, and 4) and that the RMS rises for images where there are several sharp edges (*e.g.* scenes 2 and 5). This dependence on smoothness is further demonstrated in Figures 18 and 19 which show the variation of the RMS error with the degree of local average smoothing of the original image for scene 5. The RMS error decreases with smoothing and the relative decrease is greater for the algorithm presented in this paper than for the other reconstructions.

7 Discussion

Figures 6 through 14 show the clear qualitative improvement in image resolution which is achieved by the reconstruction described in this paper and the quantitative figures of merits confirm this. It should be emphasised that, while this approach does allow a data reduction factor of a modest 33%, the primary reason for its development is not the overall image data reduction but rather that it allows a 66% reduction in the number of individual photosites on an individual sensor thereby significantly increasing the potential yield of any fabrication process.

The results presented in this paper assume perfect registration of both images and this will be an issue to be addressed when fabricating the final two-sensor camera system. However, significant problems are not anticipated as it intended to use a single lens with a beam-splitter and prism. Registration will be effected either by micropositioning of the sensors or by image translation following a calibration phase.

Perhaps the most obvious drawback of the approach is its reliance on the assumption of local smoothness in the image. In scene which exhibit strong intensity discontinuities, these assumptions are violated and the image quality does degrade somewhat in these regions. The significance of the degradation will depend on the

nature of the application, the image content, and the image processing which is used to analyse the image.

Mindful of these issues, we are presently embarking on a more general treatment of the problem. Two approaches are being pursued: the first is probabalistic, using outer products to estimate the pixel values, and the second casts the problem in linear algebraic terms and uses the Moore-Penrose pseudo-inverse to estimate the values. The power of the linear algebraic approach is that it will allow us to investigate the the exact implications of the assumptions invoked in addressing what is inherently an underconstrained problem as a function of image content (*i.e.* local grey-level pixel patterns). This work is the subject of ongoing research.

8 Conclusion

We have presented a simple technique which can effectively reconstruct a high-resolution image by combining the images generated from two sensors which exhibit an asymmetric photosite aspect ratio and asymmetric sampling density. We have done this for two configurations of the sensor geometry, one where the photosite centroids are co-linear in both directions and another where the photosite centroids are co-linear in the direction of greater sampling density (*i.e.* along the sensor columns) but the origin of each column is offset (staggered) by half the column inter-photosite distance. Good qualitative and quantitative results have been achieved with a 23% improvement in the root of sum of square error figure of merit over a simple reconstruction based on local averaging.

9 Acknowledgements

This work was funded by the European Commission as part of the Fourth Framework Programme for R&D in Information Technology (ESPRIT) under project 20557: High-Speed Manipulation by Random Access Vision (RAMAP).

10 References

1. D. Vernon, *Machine Vision*, Prenctice Hall International, 1991.
2. Philips Journal of Research, Special Issue on Solid-State Image Sensors, Vol. 48, No. 3, 1994.
3. N. Ricquier, I. Debusschere, B. Dierickx, A. Alaerts, J. Vlummens, C. Claeys, 'The CIVIS sensor: a flexible smart imager with programmable resolution', Proceeding of SPIE, Charge-Coupled Devices and Solid State Optical Sensors IV, Vol. 2172, pp. 2-10, 1994.
4. G. Meynants, B. Dierickx, D. Scheffer, A. Krymsky, 'Sensor for optical flow measurement based on differencing in space and time', Proc. SPIE Symposium on Electric Image Science & Technology, Vol. 2654, 1996.

5. D. Barry, A. Trenaman, and D. Vernon. "Random Access Sensors and their Implications for Optical Data Processing", Proceeding of the First Annual Conference of the Optical Engineering Society of Ireland, 1996.
6. G. Doemens, C. Laloni, J. Neys, and D. Scheffer. 'Random access CMOS camera with an effective resolution of 4096x4096 pixel for high speed industrial vision applications', Proceedings of the Optical Engineering Society of Ireland & Irish Machine Vision and Image Processing Joint Conference, National University of Ireland, Maynooth, pp. 1-15, 1998.
7. A. Trenaman and D. Vernon. 'Development of Generic Algorithms for Random Access Machine Vision', accepted for publication in the Proceedings of the Optical Engineering Society of Ireland & Irish Machine Vision and Image Processing Joint Conference, National University of Ireland, Maynooth, pp. 16-31, 1998.

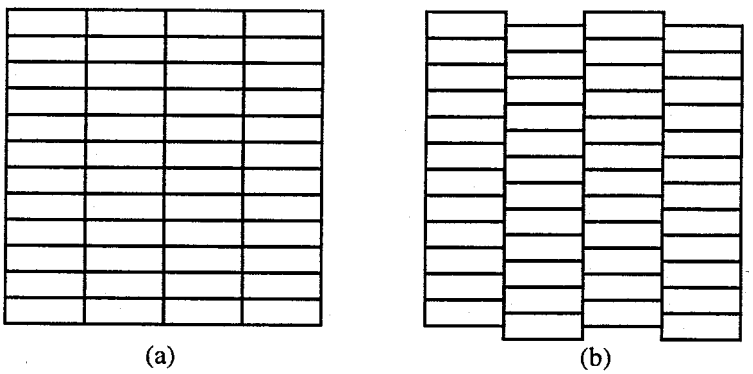


Fig. 1. 2-D sensor comprising assymetrically-shaped photosites with (a) unstaggered geometry where photosite centroids are co-linear in both directions, and (b) staggered geometry where centroids are co-linear in the direction of greater sampling density (i.e. along the sensor columns) but the origin of each column is offset by half the column inter-pixel distance. Photosites have a 3:1 aspect ratio in both cases.

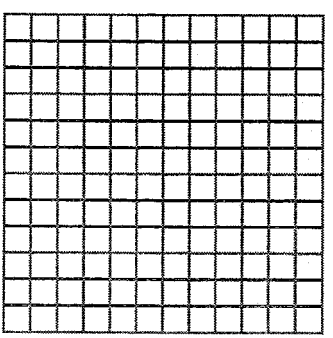


Fig. 2. Superposition of two unstaggered orthogonally-oriented images.

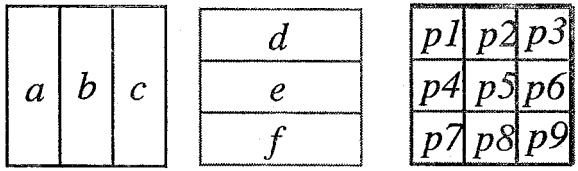


Fig. 3. Figure 3: Combination of 1×3 and 3×1 sub-regions to yield a 3×3 sub-region.

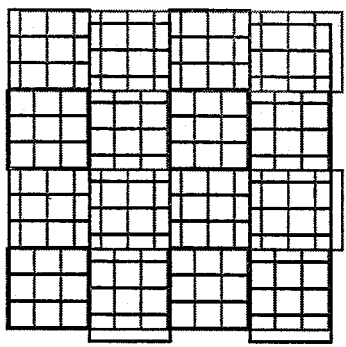


Fig. 4. Figure 4: Superposition of two staggered orthogonally-oriented images

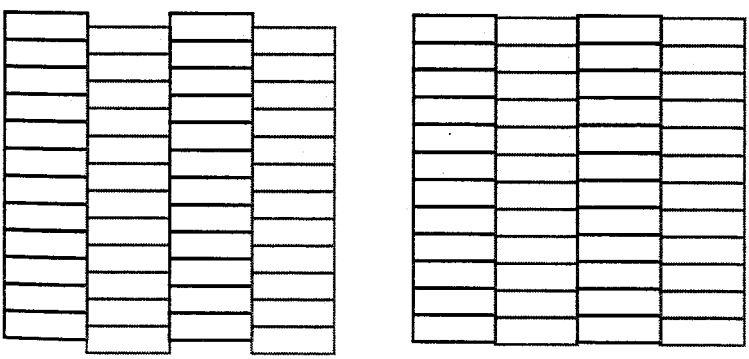


Fig. 5. Estimation of unstaggered column of pixels from staggered column.

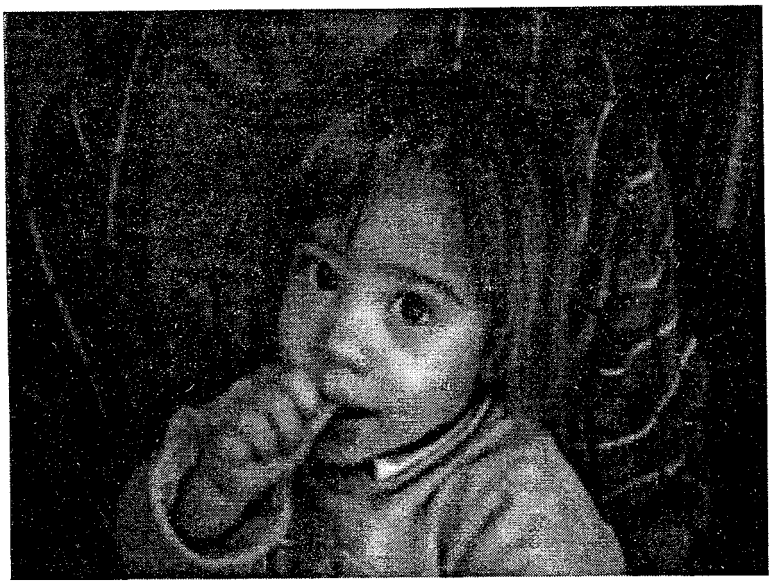


Fig. 6. Original uniformly-sampled grey-level image (scene 1: Child).

Fig. 7. Simulated image for the unstaggered geometry (rotation angle of zero degrees) with detail.

Fig. 8. Simulated image for the unstaggered geometry (rotation angle of ninety degrees) with detail.

Fig. 9. Reconstructed image using simple averaging.

Fig. 10. Reconstructed image using the model presented in this paper.

Fig. 11. Simulated image for the staggered geometry (rotation angle of zero degrees) with detail.

Fig. 12. Simulated image for the staggered geometry (rotation angle of ninety degrees) with detail.

Fig. 13. Reconstructed image using simple averaging.

Fig. 14. Reconstructed image using the model presented in this paper.

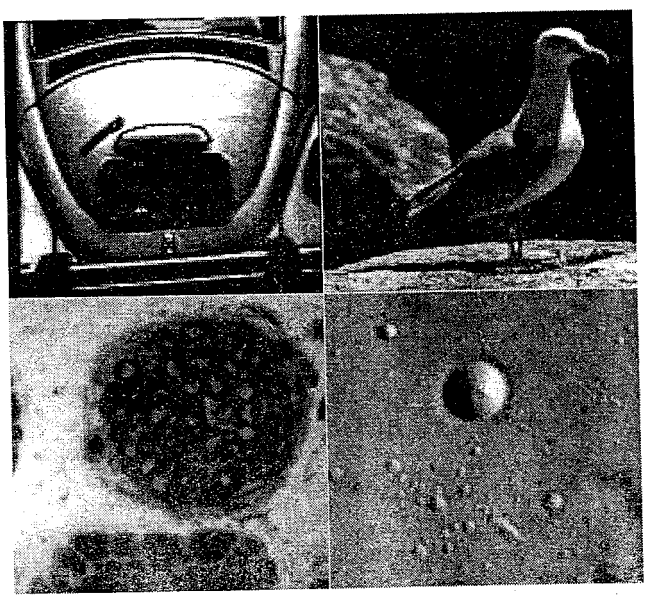


Fig. 15. Original uniformly-sampled grey-level images (top-left) Scene 2: Seagull (top-right) Scene 3: Nervecell; (bottom-left) Scene 4: Moonscape; (bottom-right) Scene 5: VW boot. Refer to Fig. 6 for Scene 1: Child.

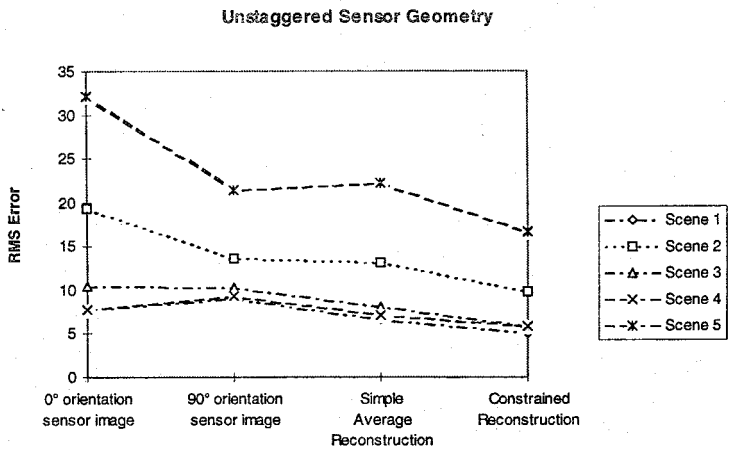


Fig. 16. Figure-of-merit values for the unstaggered sensor geometry. These figure-of-merit values are estimated by computing the root mean square (RMS) of the difference values (pixel errors) between the reconstructed image and the orginal image for five scenes.

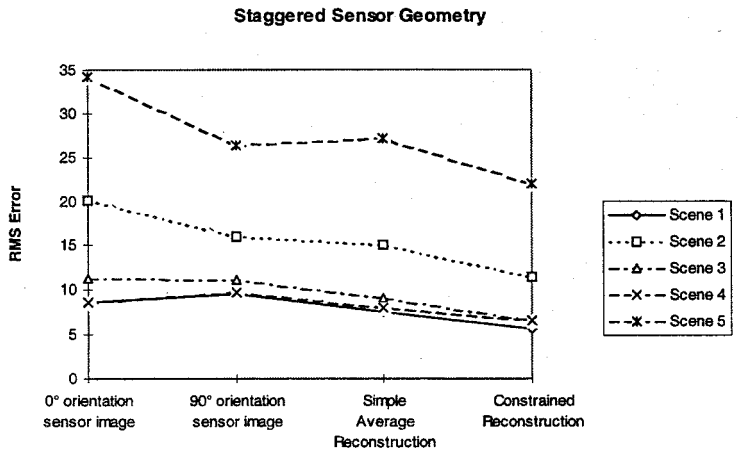


Fig. 17. Figure-of-merit values for the staggered sensor geometry. These figure-of-merit values are estimated by computing the root mean square (RMS) of the difference values (pixel errors) between the reconstructed image and the orginal image for five scenes.

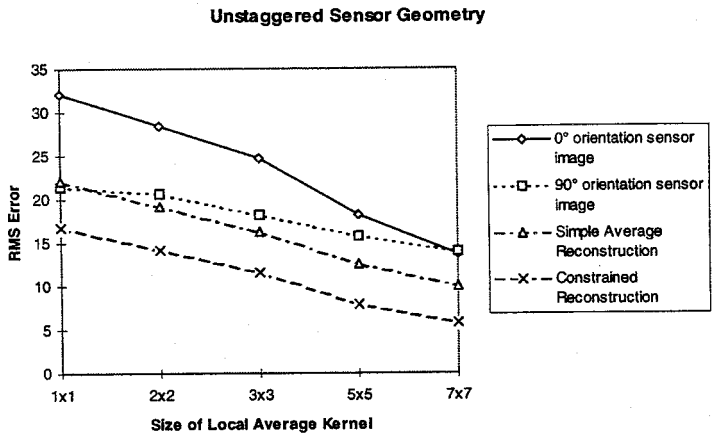


Fig. 18. Figure-of-merit values for scene number 5 (VW Boot) as a function of the degree of smoothing of the original image; unstaggered sensor geometry. As before, these figure-of-merit values are estimated by computing the root mean square (RMS) of the difference values (pixel errors) between the reconstructed image and the orginal (smoothed) image.

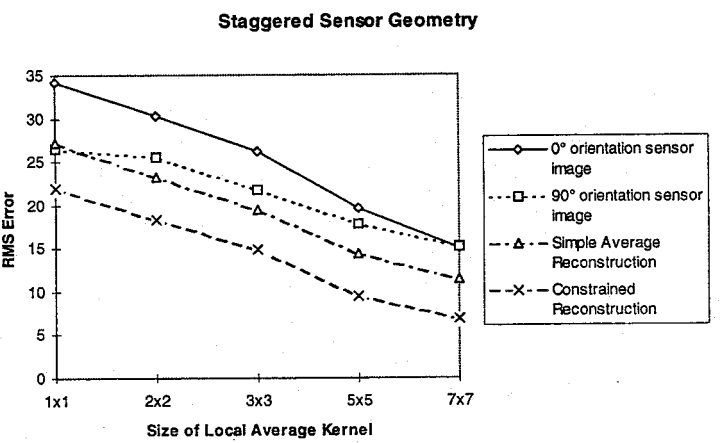


Fig. 19. Figure-of-merit values for scene number 5 (VW Boot) as a function of the degree of smoothing of the original image; staggered sensor geometry. As before, these figure-of-merit values are estimated by computing the root mean square (RMS) of the difference values (pixel errors) between the reconstructed image and the orginal (smoothed) image.

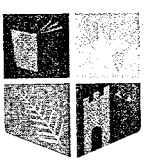
David Vernon (Ed.)

OESI–IMVIP '98

Optical Engineering Society of Ireland & Irish Machine Vision and Image Processing Joint Conference

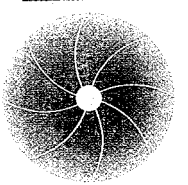
PROCEEDINGS

National University of Ireland, Maynooth
9th-10th September, 1998



NUI MAYNOOTH
Ollscoil na hÉireann Má Nuad

Cumann Innealtóireacht Optúla



THE OPTICAL ENGINEERING SOCIETY OF IRELAND