It was found that the progressive sequences showed better subjective quality than the corresponding interlaced sequences. It was estimated that the interlaced sequences needed around 50% more bits to obtain the same subjective score as the progressive sequences.

The other comparison was on subjective quality with different picture-coding formats. In our tests, the highest format always came out best. For the two sequences tested here, the MPEG-2 coding algorithm seems to work best for formats that result in quantization parameters somewhat higher than 20.

APPENDIX

Some important MPEG-specific coding parameters used in these experiments are listed below.

- Coding structure:
  - $M = 3$ for all experiments
  - $N = 12$ for interlaced sequences, i.e. sgIBBPBBPBBPg-BBIBB..., where s is sequence header, g is group header, I is intra frame, B is bidirectionally predicted frame and P is predicted frame.
  - $N = 24$ for progressive sequences
- Quantization:
  - q_scale_type = 0, which means that the value of Q is the actual divisor used. Q is in the range 2, 4, 6, . . . ,60, 62.
  - The quantization matrices from TM4 [7] were used.
- All progressive sequences were coded in MPEG-1 mode, i.e., using only frame prediction and frame DCT.
- Miscellaneous:
  - alternate_scan = 0.
  - intra_vlc_format = 1.
  - intra_DC_precision is 8 b.

REFERENCES

[1] ISO/IEC JTC 1/SC 29/WG 11, Geneva, *ISO/IEC CD 13818–2: Information technology: generic of coding moving pictures and associated audio information, Part 2: Video,* Dec. 1993.

[2] A. Puri, "Video coding using the MPEG-2 compression standard," in *Proc. SPIE's Visual Commun. Image Processing,* 1993, pp. 1701–1713.

[3] A. H. Wong and C. Chen, "A comparison of ISO MPEG-1 and MPEG-2 video coding standards," in *Proc. SPIE's Visual Commun. Image Processing,* 1993, pp. 1436–1448.

[4] CCITT SGXV Working Party SV/4, *Ref. model 8 and resulting CCITT Recommend. H.261 Video Codec Audiovisual Svcs. P*64 kb/s,* May 1989.

[5] CCIR, Geneva, *Recommend. 601-1: encoding parameters of digital television for studios.*

[6] CCIR, Geneva, *Recommend. 500-5: method for the subjective assessment of the quality of television pictures.*

[7] ISO/IEC JTC 1/SC 29/WG 11, Geneva, *MPEG93/225: test model 4,* Jan. 1993.

A Method for Rectifying Grid Junctions in Grid-Coded Images Using Cross Ratio

Tsorng-Lin Chia, Zen Chen, and Chaur-Jou Yueh

Abstract— **In the grid-coding technique, the accurate grid junction locations are vital to the surface computation. Noise arising from the imaging formation and processing processes will jeopardize the computation accuracy. The objective of this study is to make use of the geometric property of cross ratio to improve the accuracy of grid junction locations in grid-coded images. A method is devised to regenerate grid junctions from the original (i.e., raw) junction data of an image. Through a statistical error analysis, we show that the regenerated junction data are generally more accurate than the old ones in the image taken from a wide range of viewing angles.**

I. INTRODUCTION

Grid coding is a well-known technique in which two sets of parallel light planes, formed by passing an expanded beam of a collimated laser light through a grid plate, are cast on an object surface. Then the resultant grid pattern on the object surface is imaged as a projected grid pattern in the image plane. From the imaged grid pattern, the object surface can be recovered [1]–[9]. The information used for the three-dimensional (3-D) object surface computation includes the grid junction locations, grid junction slopes, and grid junction distances obtained from the image. No matter what particular technique is adopted, the accuracy of the grid junction locations is vital to the surface computation. Noise arising from the imaging formation and processing processes will jeopardize the computation accuracy.

The objective of this correspondence is to make use of the geometric property of cross ratio to improve the accuracy of junction locations in grid-coded images. We shall develop a technique based on the invariant property of cross ratio [10], [11] to use the original (i.e., raw) junction data in the image to regenerate the new junctions. We shall show, through an error analysis, that the regenerated junctions are generally more accurate than the original ones in the statistical sense. Computer experiments using both synthetic and real grid-coded images are provided to confirm that regenerated junction data are more accurate than the old junction data in the presence of noise. Besides, the junction regeneration process is valuable to the 3-D reconstruction of regular or arbitrary object surfaces.

II. REGENERATION OF PROJECTED GRID JUNCTIONS

A. The Regeneration Process

Let $G_0'', G_1'', \cdots$, and G_n'' be $n+1$ grid points on a line located on the grid plate. Assume their perspective projections are $G_0, G_1, \cdots, G_n$ on the image plane. Besides, $G_0'', G_1'', \cdots$, and G_n'' are such that $G_i'' - G_0'' = id$, where $i = 1, 2, \cdots, n$ and d is the unit grid length. Since the image points $G_0, G_1, \cdots$, and G_n are likely

Manuscript received November 4, 1994; revised November 8, 1995. This work was supported by the National Science Council of the ROC under Contract Grant NCS80-0408-E-009-10. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Moncef Gabbouj.

T.-L. Chia is with the Department of Electrical Engineering, Chung Cheng Institute of Technology, Taoyuan, Taiwan, R.O.C.

Z. Chen and C.-J. Yueh are with the Institute of Computer Science and Information Engineering, National Chiao Tung University, Hsinchu, Taiwan, R.O.C (e-mail: zchen@csie.nctu.edu.tw).

Publisher Item Identifier S 1057-7149(96)05265-7.

corrupted by noise, it is highly desirable to design a mechanism to reduce the noise in these points. We know these $n+1$ image points are constrained and cannot act totally independently. To be more specific, given any three points of them, the other points can be determined based on the cross ratio property [10], [11]. The problem is how to select the three points such that the newly regenerated image points will become more accurate. In choosing these three points, we must observe that the $n+1$ points must be collinear, as they are the image of a straight grid line lying on the grid plate. Assume that the location errors of the original $n+1$ image points are virtually identical and independent. Then the line connecting G_0 and G_n will be closer to the true (i.e., noiseless) line than any other line connecting any two different points. Therefore, we choose G_0 and G_n as two of the three wanted points. Next, which is the third wanted point? We shall represent the point relationships of $G_0, G_1, \cdots, G_n$ by a parametric form

$$G_i = G_0 + t_i(G_n - G_0) \quad \text{for } i = 0, 1, 2, \cdots, n$$

where $0 \le t_i \le 1$. Let the third wanted point be referred to as G_r with a parametric value t_r, and let G_r correspond to G_r'' on the grid plate. Then, from the cross ratio invariant property, for $i = 1, 2, \cdots, n-1$

$$\frac{t_i(1-t_r)}{(t_i-t_r)(1-0)} = \frac{id(nd-rd)}{nd(id-rd)} = \frac{i(n-r)}{n(i-r)}. \tag{1}$$

To find the optimum third point, we let an uncertain quantity Δ_r be added to t_r, say, $t_r' = t_r + \Delta_r$. Then this new t_r' will lead to a new t_i' in order to preserve the cross ratio invariance. The relation between t_r' and t_i' is given by

$$\frac{t_i'(1-t_r-\Delta_r)}{(t_i'-t_r-\Delta_r)} = \frac{t_i(1-t_r)}{t_i-t_r}.$$

After some manipulations, it yields

$$t_i' - t_i = \Delta_r \frac{t_i(t_i-1)}{t_r^2 - t_r + \Delta_r t_r - \Delta_r t_i}.$$

To minimize the effect of Δ_r on the error in $t_i' - t_i$, let $\partial(t_i' - t_i)/\partial t_r = 0$. This yields

$$t_r = (1-\Delta_r)/2 \approx 0.5 \tag{2}$$

for a very small Δ_r and $|\max(t_i' - t_i)| = |\Delta_r|$. That is, the G_r point with $t_r = 0.5$ should be chosen for the third wanted point in order to minimize the influence of the error Δ_r in t_r over the value of t_i for all i.

Now we want to find the grid cell number r associated with the corresponding point G_r'' of the point specified by $t_r = 0.5$. This can be derived from the cross ratio invariance, i.e., $\mathrm{CR}(G_0, G_i, G_r, G_n) = \mathrm{CR}(G_0'', G_i'', G_r'', G_n'')$ for a given value of i in the set $\{1, 2, \cdots, n-1\}$ together with its corresponding value of t_i. Substituting $t_r = 0.5$ into (1) yields

$$r \equiv r_i = \frac{in(1-t_i)}{i + (n-2i)t_i}, \qquad i = 1, 2, \cdots, n-1. \tag{3}$$

Here, r_i denotes the value of grid cell number r that is derived from the cross ratio equation involving G_i. Theoretically speaking, $r_1, r_2, \cdots$, and r_{n-1} are all equal. However, if there is noise in the G_i's, these values are only approximately equal. We use the average of these values as $\bar{r}$, i.e., $\bar{r}$ is equal to

$$\bar{r} = \frac{1}{n-1}\sum_{i=1}^{n-1} r_i. \tag{4}$$

Using $\bar{r}$ in (4), along with $t_r = 0.5$, and the two given endpoints G_0 and G_n, we can regenerate the new projected grid junction for t_i, denoted by t_i^{new}, from the cross ratio equation

$$\frac{t_r(1-t_i^{\text{new}})}{1(t_r - t_i^{\text{new}})} = \frac{\bar{r}(n-i)}{n(\bar{r}-i)}, \qquad i = 1, 2, \cdots, n-1$$

i.e.,

$$t_i^{\text{new}} = i(n-\bar{r})/[in + \bar{r}(n-2i)]. \tag{5}$$

This is the equation to be used in the junction regeneration process that generates new grid junctions from the original ones. An observation must be made here in order to clarify the usefulness of the regeneration process defined by (5). That is, after we average all r_i's in (4), the variance of r_i is reduced by a factor of approximately $(n-1)$. Thus, the noise is reduced in the regenerated junction. Next, we shall show that the regenerated junction locations are more accurate than the original ones through a formal error analysis.

B. Error Analysis of Regenerated Grid Junctions

Assume the original projected grid junction t_i contains a noise component e_{t_i} whose mean is zero and whose standard deviation is $\sigma_{e_{t_i}}$. Let Δt_i and Δr_i be the errors of t_i and r_i, respectively; here, let $\Delta t_i = e_{t_i}$. Then, from (3)

$$\Delta r_i = \frac{\partial r_i}{\partial t_i}\Delta t_i$$

and

$$\frac{\partial r_i}{\partial t_i} = \frac{(i^2 n - i n^2)}{[t_i(n-2i)+i]^2}.$$

Furthermore

$$E[\Delta r_i] = \frac{\partial r_i}{\partial t_i} E[\Delta t_i] = \frac{\partial r_i}{\partial t_i} E[e_{t_i}] = 0$$

The variance of r_i is given by

$$\sigma_{r_i}^2 = E[(\Delta r_i)^2] = \frac{(i^2 n - i n^2)^2}{[t_i(n-2i)+i]^4}\sigma_{e_{t_i}}^2$$

or

$$\sigma_{r_i} = \frac{n^2 i - i^2 n}{[t_i(n-2i)+i]^2}\sigma_{e_{t_i}} = K(i)\sigma_{e_{t_i}} \tag{6}$$

where $K(i) = (n^2 i - i^2 n)/[t_i(n-2i)+i]^2$.

Assume the e_{t_i}'s, $i = 1, 2, \cdots, n-1$ are identically and independently distributed (i.i.d.) and $\sigma_{e_{t_1}} = \sigma_{e_{t_2}} = \cdots = \sigma_{e_{t_{n-1}}} = \sigma_{e_t}$. Then the standard deviation of $\bar{r}$ in (4) is given by

$$\sigma_{\bar{r}}^2 = \frac{1}{(n-1)^2}\sum_{i=1}^{n-1}\sigma_{r_i}^2 = \frac{\sigma_{e_t}^2}{(n-1)^2}\sum_{i=1}^{n-1}K^2(i)$$

and

$$\sigma_{\bar{r}} = \frac{\sigma_{e_t}}{n-1}\left(\sum_{i=1}^{n-1}K^2(i)\right)^{1/2} = H\sigma_{e_t} \tag{7}$$

where

$$H = \left[\sum_{i=1}^{n-1}K^2(i)\right]^{1/2} \Big/ (n-1).$$

Next, from (5) we obtain

$$\Delta t_i^{\text{new}} = \left(\frac{\partial t_i^{\text{new}}}{\partial \bar{r}}\right)\Delta\bar{r}$$

$$\frac{\partial t_i^{\text{new}}}{\partial \bar{r}} = \frac{(ni^2 - n^2 i)}{[\bar{r}(n-2i)+in]^2}$$

and

$$E(\Delta t_i^{\text{new}}) = E(\Delta \overline{r}) = 0.$$

Then

$$\sigma_{t_i}^{2\text{new}} = E[(\Delta t_i^{\text{new}})^2] = \frac{(n^2 i - i^2 n)^2}{[\overline{r}(n-2i)+in]^4}\sigma_{\overline{r}}^2$$

and

$$\sigma_{t_i}^{\text{new}} = \frac{(n^2 i - i^2 n)H}{[\overline{r}(n-2i)+in]^2}\sigma_{e_t} = G(i)H\sigma_{e_t} \tag{8}$$

where

$$G(i) = \frac{(n^2 i - i^2 n)}{[\overline{r}(n-2i)+in]^2}.$$

From the above analysis, we have obtained the equation for noise ratio $\sigma_{t_i}^{\text{new}}/\sigma_{e_t}$ for all junctions ($i = 1, 2, \cdots, n-1$) in the image. To consider the noise in the image as a whole, we compute the average junction location noise ratio (AJLNR) as follows:

$$\text{AJLNR} = \frac{1}{n-1}\sum_{i=1}^{n-1}\sigma_{t_i}^{\text{new}}/\sigma_{e_t}.$$

AJLNR is a function of $\overline{r}$. Recall that each value of $\overline{r}$ is associated with a different perspective image of the line object. The entire AJLNR curve gives the overall noise ratio for individual images of the line object when viewed from different viewing angles. Now let us illustrate a typical AJLNR curve to see its important characteristics.

Illustrative Example: Consider a line object consisting of 11 evenly spaced points or junctions. First, we transform it into its perspective image, which is represented by a set of projected junctions $\{G_i, i = 0, 1, 2, \cdots, 10\}$. The location of each G_i can be specified by a parameter t_i with respect to the initial and final projected junctions where $t_0 = 0$ and $t_{10} = 1$. Different perspective images of the line object can be created and specified by the different t_i value of G_i for some $i \in \{1, 2, \cdots, 9\}$. Here $G_5 (i = 5)$ is used and its parameter t_5 takes a value in the interval [0.2, 0.8] with an increment step size of 0.02. Therefore, there are a total of (0.8 − 0.2)/0.02 + 1 = 31 different grid-coded images that correspond to the images as viewed from a wide range of viewing angles associated with different t_5 values. Note that the perspective image of the line object can be also specified by the value of $\overline{r}$. The relation between $\overline{r}$ and t_5 is given by (1), i.e., $\overline{r} = 10(1 - t_5)$. Fig. 1 shows the standard deviation ratios of each junction location $G_i, i = 1, 2, \cdots, 9$ for the 31 images indicated by the $\overline{r}$ values. The AJLNR curve for all these images is also shown. The characteristics of the AJLNR curve include: i) The AJLNR value is smaller than one for the given range of $\overline{r}$ and reaches its minimum at $\overline{r} = 5$ (or $t_5 = 0.5$); ii) the curve is symmetrical with respect to $\overline{r} = 5$ (or $t_5 = 0.5$). Therefore, according to this theoretical analysis, the regenerated grid junctions are more accurate than the old ones. This will be confirmed by Experiment 1 described below.

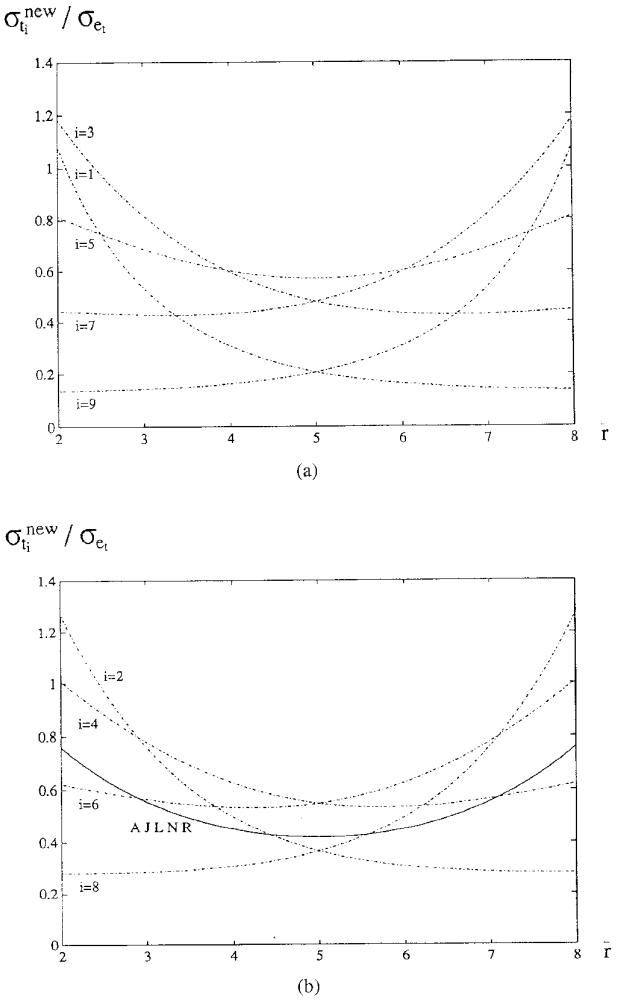


Fig. 1. Plots of the standard deviation ratios of the regenerated junction location to the original junction location for all junctions, $i = 1, 2, \cdots, 9$, and the average of the ratios for the values of $\overline{r}$ from 2.0 to 8.0.

III. Computer Simulations

A. Experiment 1

In order to examine the theoretical error model of our method, we conduct an experiment using the same line object data given in the above illustrative example. In this experiment, the noise of each projected junction location in the image is represented by a Gaussian distribution with mean = 0 and standard deviation = 0.005. One thousand noisy copies of each ideal image are generated by computer. Table I shows the case of a typical image with $r = 6$. The column with heading $t_i^{\text{ideal}}, i = 0, 1, 2, \cdots, 10$, contains the ideal junction locations; the column with heading t_i lists the actual junction locations that have been corrupted by Gaussian noise. The initial and final junctions $t_0 = 0.0005$ and $t_{10} = 0.99875$ are used as G_0 and G_n in the junction regeneration process. The normalized junction set is given in the column with heading t_i'. The reference junction r_i corresponding to the projected grid junction t_r =0.5 derived from each given t_i' is shown in the column with heading r_i. The average value of all r_i's is found to be $\overline{r}$ =5.993128. The new projected grid junctions derived from the regeneration process are given in the column with heading $t_i'^{\text{new}}$. The regenerated projected grid junctions are transferred back to the original scale by using $t_i^{\text{new}} = t_i'^{\text{new}}(0.99875 - 0.0005) + 0.0005$. The result is given in the column with heading t_i^{new}.

We then compute the standard deviation noise ratio figures directly from the original and regenerated junction data. From Table I, the average of absolute location errors of the nine original junction locations, given by

$$\frac{1}{9}\sum_{i=1}^{9}|t_i - t_i^{\text{ideal}}|$$

TABLE I
THE RESULT OF A REGENERATION PROCESS FOR THE GRID JUNCTIONS IN A TYPICAL GRID-CODED IMAGE

i	t_i^{ideal}	t_i	t_i'	r_i	$t_i'^{new}$	t_i^{new}
0	0	0.0005	0		0	0.0005
1	0.068966	0.067966	0.067584	6.051998	0.069150	0.695285
2	0.142857	0.144607	0.144360	5.970644	0.143208	0.143457
3	0.222222	0.226347	0.226243	5.944400	0.222717	0.222828
4	0.307692	0.312817	0.312864	5.941857	0.308302	0.308263
5	0.4	0.404625	0.404833	5.951665	0.400687	0.400486
6	0.5	0.498625	0.498998	6.009613	0.500716	0.500339
7	0.608696	0.604196	0.604754	6.039577	0.609377	0.608811
8	0.727273	0.726773	0.727546	5.996689	0.727840	0.727066
9	0.857143	0.854518	0.855515	6.031710	0.857493	0.856492
10	1	0.998750	1		1	0.998750

is 0.002847 and the average of absolute location errors of the regenerated junction locations, given by

$$\frac{1}{9}\sum_{i=1}^{9}|t_i^{\text{new}} - t_i^{\text{ideal}}|$$

is 0.000459. Therefore, the noise in the original junction locations is reduced by a factor of 6.2 after the application of our junction regeneration process. Furthermore, for the 1000 generated copies of each grid-coded image we computed the average of junction noise reduction figures, AJLNR, over 1000 copies. The results for the 31 grid-coded images are given in Fig 2. This figure shows that the measured and theoretical AJLNR values are in a close agreement. Thus, we confirm the theory of our theoretical error analysis presented in the previous section.

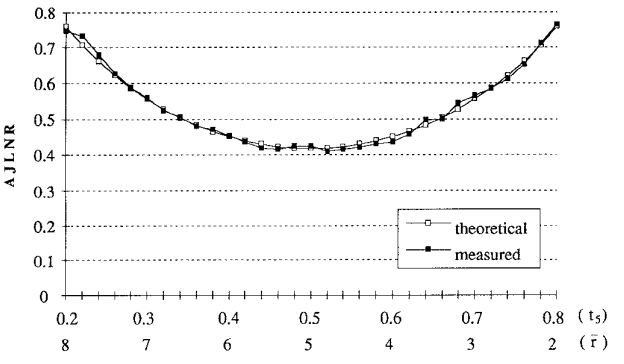
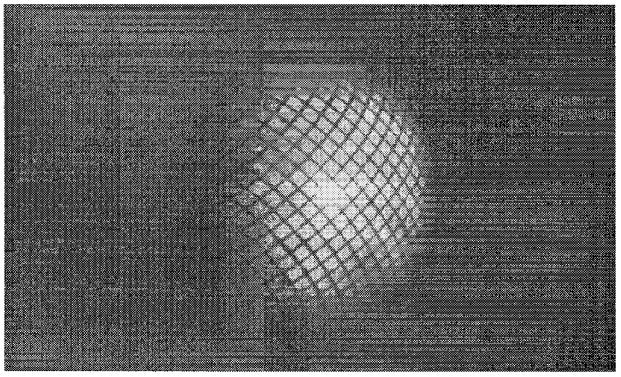


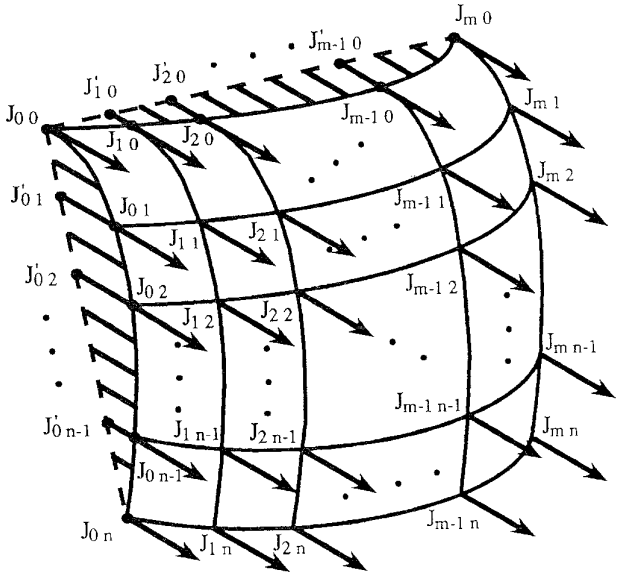
Fig. 2. Measured and theoretical values of the average junction location noise ratios for the 31 grid-coded images.

B. Experiment 2

In this experiment, we use a real grid-coded image of a typical sphere as shown in Fig. 3(a). A surface reconstruction method was reported in [12] for determining the location and radius of a sphere in 3-D space using the extracted grid junctions from its single grid-coded image. Here, we want to show the accuracy improvement of the reconstruction result by using our junction regeneration data in the presence of noise. The extracted grid junctions contain noise due to quantization and other error sources, so the sphere radius estimations with and without the junction rectification process are 51.72 mm and 46.21 mm, respectively. (The actual radius is 50 mm.) Thus, the regenerated junctions produce a better estimation result. In order to simulate more sphere images with noise, we add three sets of Gaussian noise with mean of zero and standard deviation of 1.5 pixels to the extracted grid junctions. Before we explain how to apply the cross ratio invariant property to the points on the curved stripe, we need to introduce the concept of the virtual linear grid line. In Fig. 3(b) each curved stripe—for instance, the one containing surface grid junctions $J_{0,0}, J_{0,1}, \cdots$, and $J_{0,n}$—is the intersection of a sheet of light plane with the object surface. And this light plane can be represented by the plane containing the curved stripe and a linear line connecting any two grid junctions, say, the two end junctions, $J_{0,0}$ and $J_{0,n}$, called a *virtual linear grid line*. Furthermore, all surface grid junctions are passed through by the parallel light rays, which are in turn formed by the intersection of the two sets of parallel light planes. The extensions of the parallel light rays through the surface grid junctions will intersect the virtual linear grid line at virtual grid junctions, e.g., $J'_{0,1}, J'_{0,2}, \cdots$, and $J'_{0,(n-1)}$. Therefore, we can apply the constraint of cross ratio invariance to the imaged virtual



(a)



(b)

Fig. 3. (a) Grid-coded image of a sphere of radius 50 mm. (b) Grid junctions on the curved grid lines and derived grid junctions on the virtual linear grid lines.

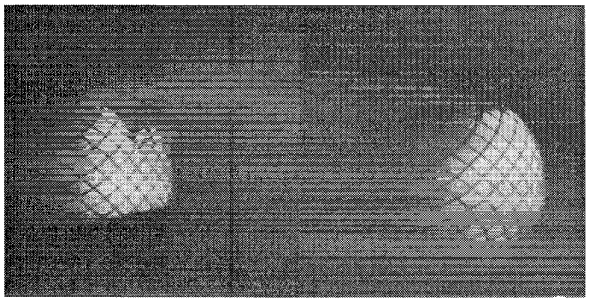
grid junctions for junction rectification. Next, each line connecting a rectified grid junction and the known vanishing point of the parallel light rays intersects the fitted imaged curved stripe at a new surface grid junction. These new grid junctions can replace the original ones for the 3-D spherical surface reconstruction. In Table II we show the computation results of the sphere radius with and without the rectification of the grid junctions. From this table, it can be seen that our regeneration process can cope with Gaussian noise and, thus, aids the 3-D surface reconstruction task.

C. Experiment 3

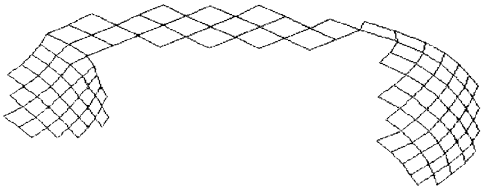
In this experiment, we want to show that our method works for any smooth curved surfaces, not just regular surfaces such as planar or spherical ones. Here we use our junction regeneration process in a grid-coded image of a telephone receiver, as shown in Fig. 4(a). By the same token described in the above experiment, the grid junctions $J_{0,0}, J_{0,1}, \cdots$, and $J_{0,n}$ will lead to the virtual grid junctions $J'_{0,1}, J'_{0,2}, \cdots$, on the virtual linear grid line and $J'_{0,(n-1)}$ and the cross ratio invariant property is used. Finally, the new surface

TABLE II
THE SPHERE RADIUS ESTIMATION WITH OR WITHOUT THE RECTIFICATION OF THE EXTRACTED GRID JUNCTIONS IN THE PRESENCE OF A GAUSSIAN NOISE WITH ZERO MEAN AND STANDARD DEVIATION OF 1.5 PIXELS (ACTUAL SPHERE RADIUS = 50 MM)

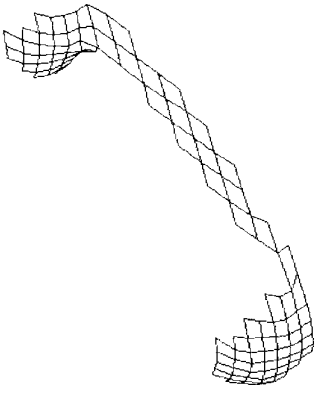
Data Set	With the Junction Rectification	Without the Junction Rectification
1	51.29 mm	46.63 mm
2	47.74 mm	43.29 mm
3	49.94 mm	45.32 mm



(a)



(b)



(c)

Fig. 4. (a) Grid-coded image of a telephone receiver. (b) View of the reconstructed 3-D visible telephone receiver surface. (c) New view of the reconstructed 3-D visible telephone receiver surface.

grid junctions are obtained and used to find the 3-D telephone receiver surface.

Fig. 4(b) shows the reconstructed surface of the telephone receiver visible in the image. Fig. 4(c) is another perspective of the receiver surface when viewed from a different viewing angle.

IV. Conclusion

The cross ratio is an important invariant property in projective geometry. It can be applied to grid-coded images. Based on this invariant property, a junction regeneration process is proposed to reproduce the grid junctions from the original junctions. From a theoretical error analysis, we show the regenerated junction data are more accurate than the old junction data in the presence of noise. The experiments are conducted and the experimental results confirm our finding. The junction regeneration process can aid in the reconstruction of regular or arbitrary object surfaces. We are currently using this junction regeneration method to construct the solid model of the objects for the CAD/CAM applications.

References

[1] P. M. Will and K. S. Pennington, "Grid coding: a novel technique for image processing," *Proc. IEEE,* vol. 60, no. 6, pp. 669–680, June 1972.

[2] Y. F. Wang, A. Mitiche, and J. K. Aggarwal, "Computation of surface orientation and structure of objects using grid coding," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. PAMI-9, no. 1, pp. 129–137, Jan. 1987.

[3] N. Shrikhande and G. Stockman, "Surface orientation from a projected grid," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. 11, no. 6, pp. 650–655, June 1989.

[4] G. Hu and G. Stockman, "3-D surface solution using structured light and constraint propagation," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. 11, no. 4, pp. 390–402, Apr. 1989.

[5] Y. F. Wang and J. K. Aggarwal, "Integration of active and passive sensing techniques for representing 3-D object," *IEEE Trans. Robot. Automat.,* vol. 5, no. 5, pp. 460–471, Dec. 1989.

[6] M. Asada, H. Ichikawa, and S. Tsuji, "Determining surface orientation by projecting a stripe pattern," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. 10, no. 5, pp. 749–752, Sept. 1988.

[7] D. C. Tseng and Z. Chen, "Computing location and orientation of polyhedral surface using a laser-based vision system," *IEEE Trans. Robot. Automat.,* vol. 7, no. 6, pp. 842–848, Dec. 1991.

[8] Z. Chen, S. Y. Ho, and D. C. Tseng, "Polyhedral face reconstruction and modeling from a single image with structured light," *IEEE Trans. Syst., Man, Cybernet.,* vol 23, no. 3, pp. 864-872, 1993.

[9] Z. Chen, T. L. Chia, and S. Y. Ho, "Measuring 3-D location and shape parameters of cylinders by a spatial encoding technique," *IEEE Trans. Robot. Automat.,* vol. 10, no. 5, pp. 632–647, 1994.

[10] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis.* New York: Wiley, 1973.

[11] G. Lei, "Recognition of planar object in 3-D space from single perspective views using cross ratio," *IEEE Trans. Robot. Automat.,* vol. 6, no. 4, pp. 432–437, Aug. 1990.

[12] T. L. Chia, "Computer vision for measuring 3-D location and shape parameters of primitive objects by structured light," Ph.D. dissertation, Chiao Tung University, Taiwan, R.O.C., 1993.

Design of Lapped Orthogonal Transforms

Richard Heusdens

***Abstract*—In this correspondence, we present some properties and new results of paraunitary filter banks. We concentrate on the case where the filter length $L = 2K$, where K is the number of channels. The aim is to design perceptually relevant filters, i.e., linear-phase filters that smoothly decay to zero at the boundaries.**

I. Introduction

Transform coding is one of the most efficient methods for data compression of correlated signals [1]–[3]. In traditional transform coding, nonoverlapped orthogonal transforms are commonly used. This is done by mapping the input sequence x into an output sequence u using an orthogonal, block-diagonal Toeplitz operator, say T. In practical situations, the sequence u is quantized, coded, transmitted, and/or stored and reconstructed before being inversely mapped. The resulting sequence y will, therefore, in general be an approximation of the input sequence x. In case $y = x$ we say that y is a perfect reconstruction (PR) of the input sequence x. A popular example of a nonoverlapped orthogonal transform is the discrete cosine transform (DCT) [1]–[4]. It is well known that the disadvantage of nonoverlapped transforms is that "blocking artifacts" are introduced at low bit rates. This can be seen as follows. The effect of quantization can be regarded as adding an additional noise signal to the sequence u. Hence, the reconstruction y becomes $y = T^t(u + n) = x + \varepsilon$, i.e., y can be written as x, a perfect replica of the input signal, and an error term ε, which is as a linear combination of the columns of T^t. We shall refer to these columns as the basis functions of the transform. In case of nonoverlapped transforms, the basis functions have abrupt changes at the endpoints of their support, which causes equally spaced discontinuities in the error term. To avoid these artifacts, we should choose basis functions without abrupt changes that tend to zero smoothly at the endpoints of their supports. However, this is not possible with nonoverlapped transforms when the PR property is to be maintained at the same time.

A class of transforms that can eliminate the blocking artifacts is the class of so-called 50% overlapped transforms [5]–[7]. In this case, the map T has a band-diagonal Toeplitz structure, where the nonzero parts of the basis functions overlap one another by 50%. The orthogonal transforms of this subclass are collectively referred to as the *lapped orthogonal transform* (LOT). In [5]–[7], it was shown that the LOT can be constructed by using a DCT. However, the basis functions thus obtained still have a certain amount of discontinuity, which will also lead to blocking artifacts. In [8], a solution is proposed for overcoming this problem. However, this solution leads to nonorthogonal transforms. In this paper we will design LOT's that do not introduce blocking artifacts.

In [5], it was shown that transform coding can be regarded as a special case of multirate filter bank coding. Placing transform coding into this more general framework has the advantage that the basis functions can be viewed as the impulse responses of the bank filters, downsampled to their Nyquist rate. Therefore, all properties known

Manuscript received April 26, 1994; revised November 20, 1995. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Dan Schonfeld.

The author is with Philips Research Laboratories, 5656 AA Eindhoven, The Netherlands (e-mail: heusdens@natlab.research.philips.com).

Publisher Item Identifier S 1057-7149(96)05260-8.