Toward Optimal Structured Light Patterns

Eli Horn and Nahum Kiryati
Department of Electrical Engineering
Technion - Israel Institute of Technology
Haifa 32000, Israel

Abstract

A methodology for the optimal design of projection patterns for stereometric structured light systems is presented. It draws on the similarity as well as the difference between the design of projection patterns and the design of optimal signals for digital communication. Seemingly unrelated structured light methods, such as the Gray code scheme and intensity ratio techniques, are unified as special cases within the suggested theoretical framework. The design of K projection patterns for a structured light system with L distinct planes of light turns out to be equivalent to the placement of L points in a K dimensional space subject to certain constraints. Optimal design in the MSE sense can be defined, but leads to an intractable multi-parameter global optimization problem. Intuitively appealing suboptimal solutions are derived from the family of K dimensional space-filling Hilbert curves. The theoretical results are supported by experiments.

Keywords: code space, intensity ratio, range imaging, space filling curve, structured light

1 Introduction

A structured light stereometric system [3] is similar to a passive stereo vision system with one of the cameras replaced by a projector. In its conceptually simplest form a laser projector illuminates a single point in the scene. Assuming that the illuminated point is visible by the camera, it can be easily identified in the image. This is the great advantage of structured light techniques over passive stereo vision in which the fundamentally difficult correspondence problem must be solved. The intersection of the known illumination ray and the line of sight defines the 3-D location of the illuminated point in the scene. For dense reconstruction the scene must be scanned by the laser beam, a time consuming procedure.

The acquisition time can be reduced by using stripe illumination. A light source, normally a laser with a cylindrical lens, projects a plane of light that creates a narrow stripe on the scene [1, 19].

Since the intersection of a known illumination plane and a line of sight uniquely determines a point, the 3-D location of all points along the stripe that are visible by the camera can be simultaneously obtained. For dense reconstruction the scene must be scanned by the plane of light, still a fairly slow process. Specialized sensors [2, 11] are a faster alternative to standard CCD cameras in this application.

In order to further speed up the range sensing process and eliminate mechanical scanning and laser illumination, various systems that use a spatially modulated white light projector have been suggested. In such systems many light planes or rays can be projected simultaneously as parts of a single illumination pattern. Then, however, their correct identification in the image is a crucial problem that must be addressed.

Several single illumination pattern systems in which light planes or rays are identified by certain markings have been suggested in the literature, e.g. [5, 12, 14, 22]. Since the markings occupy some space that must remain connected in the image of the illuminated scene, the applicability of these techniques is generally limited to "well-behaved" scenes. Blake *et al* [4] have been able to overcome this limitation, but require a carefully aligned *trinocular* system with two projectors and a camera. In the rainbow range-finder [21, 20] a white illumination source is dispersed into a beam of monochromatic light-planes that can, in principle, be uniquely identified by an appropriate color camera.

High reliability identification of light planes with minimal assumptions on the nature of the scene can be achieved by *spatio-temporal* modulation of the illumination, i.e., by sequentially projecting several patterns [15]. A robust and widely applied structured light system that is based on spatio-temporal modulation has been described in [17]. Gray code is used to label the light planes, and each illumination pattern represents one of the bit planes of the Gray code labels. Consider a given

location in the image plane. The plane of light that impinges upon the corresponding scene point can be uniquely identified: its Gray code label is obtained by noting under which illumination patterns that point is lit. The number of projection patterns needed is in principle equal to the number of bits required for light plane labeling at the desired resolution. Since the power of the light source is distributed across much of the scene, local image brightness is much smaller than in a single ray or a single plane system. In order to account for variability in the reflectivity of surfaces in the scene, spatially varying bit-decision thresholds must be carefully set. They can be derived using two additional images, one taken under full uniform illumination and the other under just the ambient light. Alternatively, the negative as well as the positive illumination patterns that correspond to each bit plane can be projected (thus doubling the number of patterns), allowing threshold-free binary decisions to be made according to the local sign of the difference between the positive and negative images.

In the intensity ratio depth sensor [6] grey levels are used for light-plane labeling. Two illumination patterns are required. One is a linear intensity wedge that varies horizontally from half to full intensity. The other illumination is uniform. The image taken under wedge illumination is divided pixel-wise by the reference image that was obtained under uniform illumination. Assuming that the brightness of any point in the scene is proportional to the intensity of the light plane on which it lies, the value at any point in the ratio image identifies the corresponding light-plane. This system is quite sensitive to noise.

Chazan and Kiryati [8] developed a *pyramidal* intensity ratio scheme. They proposed to employ the result of the basic single wedge intensity ratio depth sensor as a first approximation. A double wedge is then used in the second stage, a quadruple wedge in the third stage etc., thus progressively increasing the slopes of the wedges and improving the accuracy. From the second stage on, for each intensity ratio value there are several possible originating planes of light. The correct one is identified

according to the approximate results obtained in the previous stage. Since reliable information cannot be extracted near grey level discontinuities, illumination patterns shifted by half wedge-length are also used.

It is interesting to compare the Gray code system [17] to the pyramidal intensity ratio depth sensor [8]. Both techniques use spatio-temporal modulation of the illumination source. Furthermore, both methods follow the coarse-to-fine paradigm. This means that the ability to resolve adjacent light planes, hence depth measurement accuracy, improves as additional illumination patterns are added. The Gray code bit plane illumination patterns could be regarded as the quantization to a single bit of some continuous functions. Since one or two additional reference images are taken under uniform illumination to derive spatially varying bit-decision thresholds, the Gray code method can be regarded as a degenerate hierarchical intensity ratio system. Observe that in the Gray code scheme each illumination pattern provides just a single bit to the light plane labels. Thus, at low noise levels, the initial depth measurement accuracy of the intensity ratio range sensor, and the accuracy obtainable after a few additional stages of pyramidal improvement, are greater than in the Gray code technique with the same number of illumination patterns. However, the large noise margins in the Gray code method make it more robust to higher noise levels.

With spatio-temporal modulation of the light source, robust identification of light planes can be achieved with minimal assumptions on the smoothness of the scene. However, the applicability of range sensing systems of this type is limited to scenes that are static within the acquisition period. Since an image of the scene must be taken under each illumination pattern, the range acquisition period is lower bounded by the product of the number of illumination pattern and the time needed for acquiring a single image of the scene. Because the latter is normally dictated by technology (video rate), the number of projection patterns must be as small as possible subject to the application-specific accuracy requirements.

The following fundamental problem thus arises in the design of a structured light system that relies on spatio-temporal modulation of the light source: Given the total number of projection patterns that can be used and the noise level in the system, what are the projection patterns that lead to the best accuracy that can be achieved. Alternatively, given the noise level in the system, design the smallest set of projection patterns that meets the application specific accuracy requirements. An approach to solving this problem is outlined in this paper.

2 The Optimal Design Problem

In this paper the problem of optimal design of spatio-temporal modulation schemes for structured light systems is addressed. We begin by observing the similarity of a structured light system to a digital communication system. The projector transmits a unique temporal code for each of the L light planes. At each point in the image plane, a noisy transmission is received and subsequently decoded. This process is carried out in parallel for all points in the image plane. A K bit bilevel code is used in the Gray code scheme, hence up to 2^K planes of light can in principle be distinguished. A scalar code with a large alphabet of intensity ratios is employed in the basic intensity ratio range sensor [6]. The pyramidal intensity ratio system [8] is similar, but there the code is a vector of intensity ratios.

When K illumination patterns are used (in addition to a uniform illumination reference used for normalization), the horizontal index x of each plane of light is encoded by a vector $\vec{f}(x)$ of length K that can be represented as a point in a K dimensional space that is referred to as the *code space*. The collection of code words that are attached to all possible horizontal indices is a set of points $F = \{\vec{f}(x) | x \in X\}$ in the code space. This set of transmitted points F can be regarded as a digital curve in the code space, ordered by x. Since the allocation of code words to planes of light should

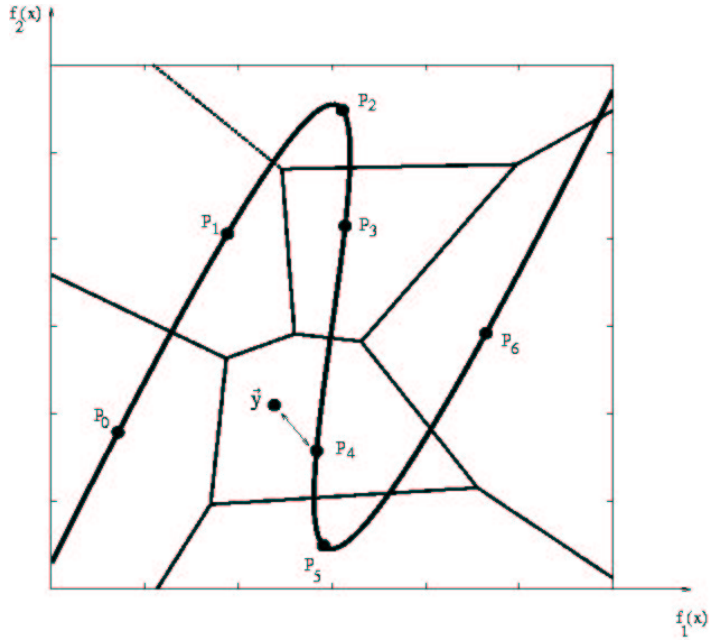

Figure 1: The code space representation of a two-dimensional ($K = 2$) coding scheme with $L = 7$ code words. The digital curve point that is nearest to the noisy vector of measurements $\vec{y}$ is P_4.

be unique, the digital curve must not self-intersect.

The vector $\vec{y}$ of K normalized values measured at a certain observed location in the scene under the K illumination patterns can be represented by a point in the K dimensional code space. Ideally, the point would coincide with one of the digital curve points and thus identify the horizontal index x of the illumination plane that impinged upon that scene location. In practice, due to noise, this does not happen. The digital curve point that is nearest (in the L_2 sense) to the noisy measurement vector is taken as an estimate for the original light plane:

$$\hat{x} = \mathrm{NS}(\vec{y}) \equiv \arg\min_{x} ||\vec{y} - \vec{f}(x)||_2 \tag{1}$$

For illustration, Fig. 1 shows the code space representation of a two-dimensional ($K = 2$) coding scheme with $L = 7$ code words. The digital curve point $\mathrm{NS}(\vec{y})$ that is nearest to the noisy vector of measurements $\vec{y}$ is P_4.

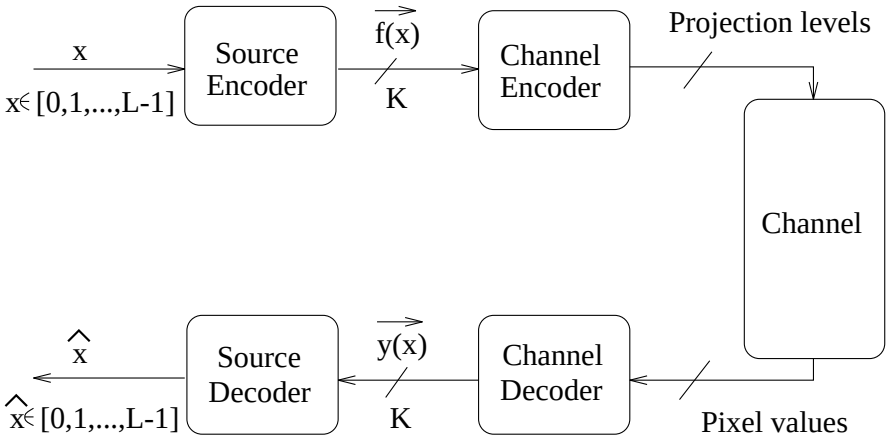

Figure 2: The analogy between a structured light system and a digital communication system.

The analogy between a structured light system and a digital communication system is summarized in Fig. 2. Given an index x of a light plane, the "source encoder" generates a code word $\vec{f}(x)$. The "channel encoder" represents $\vec{f}(x)$ by the projection brightness levels of that light plane. The output of the "channel" is a set of brightness readings in a given pixel. The "channel decoder" transforms these readings to a vector of normalized measurements $\vec{y}$. Finally, the estimate $\hat{x}$ of the index x is computed in the "source decoder".

In schemes based on intensity ratios, the received brightness values are normalized. This is done by pixel-wise division with a reference image obtained under uniform illumination. The derivation of spatially varying bit decision levels in the Gray code scheme can be regarded as a special case of this procedure. The channel encoders and decoders are thus similar in these methods, and the difference is primarily in the source encoders and decoders. Fig. 3 shows the Gray code scheme and the pyramidal intensity ratio technique within this framework. The channel encoding and decoding are omitted for clarity.

In a digital communication system, the primary quality indicator is the probability of error. If an erroneous code word is received, the cost is usually independent of its identity. In contrast, in a

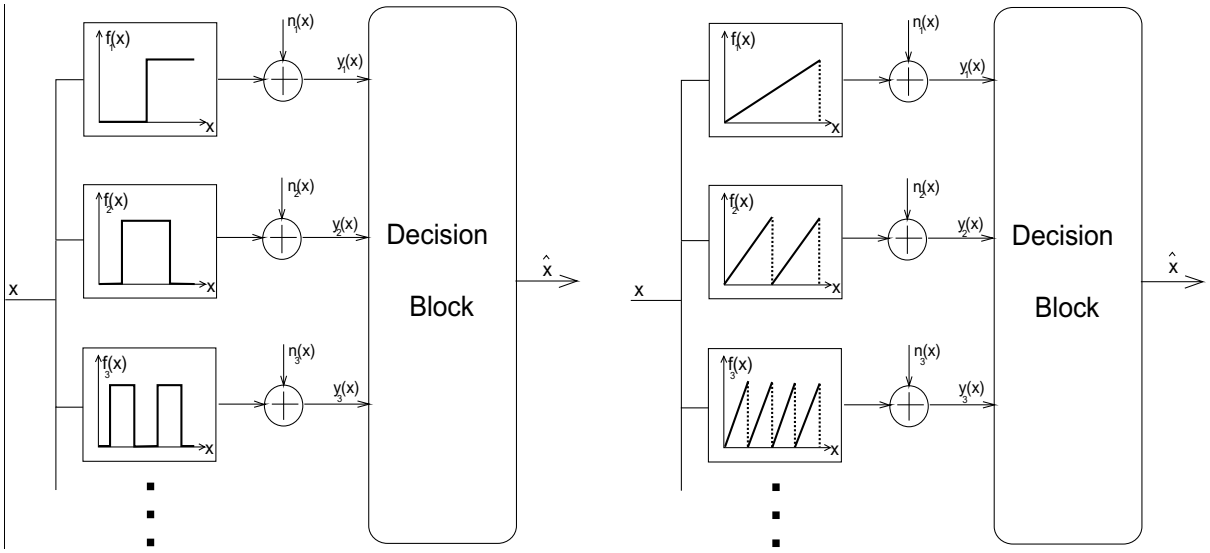

Figure 3: The Gray code (left) and the pyramidal intensity ratio (right) structured light systems viewed as digital communication systems.

structured light system each depth measurement is subject to some error, and the figure of merit of the system must be related to the magnitude of the depth errors. The depth error depends on the difference between the true index x of the plane of light and the estimated index $\hat{x}$. The exact relation is quite complex, but normally the size of the depth error is roughly proportional to $|x - \hat{x}|/L$. We proceed to design coding schemes that give the best MSE estimates of the transmitted planes of light, i.e., that minimize

$$\varepsilon_x^2 = E_{x,\vec{y}} \left[\frac{x - \hat{x}}{L}\right]^2 . \tag{2}$$

The averaging is over all the possible indices of planes of light that can be transmitted and over all the possible measurement vectors $\vec{y}$ that can be received.

Generally,

$$\varepsilon_x^2 = \sum_{x=0}^{L-1} \int_{\vec{y}} \left[\frac{x - \hat{x}}{L}\right]^2 \cdot p(\vec{y}|\vec{f}(x))\mathrm{Prob}(\vec{f}(x))d\vec{y} . \tag{3}$$

Suppose that all the noise and error sources in the system could be represented by an additive random

noise vector $\vec{n}$ that is independent of the light plane:

$$\vec{y} = \vec{f}(x) + \vec{n} \ . \tag{4}$$

Assume also that the a-priori probability of appearance of the planes of light is identical, so

$$\text{Prob}(\vec{f}(x)) \approx \frac{1}{L} \quad 0 \leq x \leq L-1 \ . \tag{5}$$

Under these simplifying assumptions and taking $\hat{x} = \text{NS}(\vec{y})$, ε_x^2 can be approximated by

$$\varepsilon_x^2 \approx \frac{1}{L} \sum_{x=0}^{L-1} \int_{\vec{y}} \left[\frac{x - \text{NS}(\vec{y})}{L} \right]^2 \cdot p_{\vec{n}}(\vec{y} - \vec{f}(x)) d\vec{y} \ . \tag{6}$$

Note that if it can be assumed that the density function of the noise vector is monotonically decreasing with the L_2 norm $||\vec{n}||_2$, i.e.,

$$||\vec{n_1}||_2 > ||\vec{n_2}||_2 \quad \longrightarrow \quad p_{\vec{n}}(\vec{n_1}) < p_{\vec{n}}(\vec{n_2}) \quad \forall \ \vec{n_1}, \vec{n_2} \ , \tag{7}$$

then $\text{NS}(\vec{y})$ is the *maximum likelihood* estimator for the horizontal index x of the original light plane:

$$\hat{x}_{ML} = \arg\max_x p_{\vec{n}}(\vec{y}|\vec{f}(x)) = \arg\max_x p(\vec{y} - \vec{f}(x)) = \arg\min_x ||\vec{y} - \vec{f}(x))||_2 \equiv \text{NS}(\vec{y}) \ . \tag{8}$$

Theoretically, the MSE estimator $\hat{x}_{MSE} = E(x|\vec{y})$ is a better choice than $\hat{x}_{ML} = \text{NS}(\vec{y})$ for the minimization of ε_x^2 in Eq. 2. It can be expressed as

$$\hat{x}_{MSE} = E(x|\vec{y}) = \sum_{x=0}^{L-1} x \cdot \text{Prob}(x|\vec{y}) = \frac{\sum_{x=0}^{L-1} x \cdot p(\vec{y}|x)\text{Prob}(x)}{\sum_{x=0}^{L-1} \cdot p(\vec{y}|x)\text{Prob}(x)} \tag{9}$$

$\hat{x}_{ML}$ is however considerably easier than $\hat{x}_{MSE}$ to compute. Experiments [10] have shown that the added accuracy gained by applying $\hat{x}_{MSE}$ rather than $\hat{x}_{ML}$ is insignificant. Therefore $\hat{x}_{ML}$ is used throughout this research.

Our goal is to design a set of code words $\{f(x)|x = 0, \ldots, L-1\}$, that minimizes ε_x^2. From Eq. 6 it is obvious that the optimal solution depends on the noise distribution. In practice, the noise in the system is the combined outcome of various sources. We assume that the overall noise distribution can be approximated by a zero-mean Gaussian with a given standard deviation. Finding the coding scheme that minimizes ε_x^2, even in the simplified Eq. 6, is a formidable global optimization problem with $L \cdot K$ parameters. For example, with $L = 512$ light planes and $K = 5$ projection patterns, the global optimization should take place in a 2560-dimensional space, clearly a difficult task.

3 Suboptimal Design

The design of K projection patterns for a structured light system with L light planes is equivalent to the placement of L points $\{f(x)|x = 0, \ldots, L-1\}$ in a K dimensional space. It is desirable to keep each point as far as possible from all the others. But since with the MSE criterion the error penalty is proportional to $(x - NS(\vec{y}))^2$, it is most important to keep two points $f(x_1)$, $f(x_2)$ far from each other if $|x_1 - x_2|$ is large. Where $|x_1 - x_2|$ is small, keeping a large distance between $f(x_1)$ and $f(x_2)$ is less important.

In terms of the representation of the code words $\{f(x)|x = 0, \ldots, L-1\}$ as a digital curve ordered by x, the main requirement is that any two code points that are distant from each other along the curve should be sufficiently distant in the code space. This section focuses on the development of suboptimal, intuitively appealing code space arrangements derived from families of space filling curves such as the family of *Hilbert* space filling curves [16].

The n-dimensional Hilbert curve is a transformation from the segment $[0,1]$ to the hypercube $[0,1]^n$. Let $T_{h_{2D}}$ denote the 2-dimensional Hilbert curve:

$$T_{h_{2D}} = \begin{pmatrix} f_1 \\ \\ f_2 \end{pmatrix} (t) \ : \ [0,1] \stackrel{onto}{\longrightarrow} [0,1]^2 \tag{10}$$

Let t be represented in base-4 as $t = 0.t_1 t_2 t_3 \ldots t_n$. Then

$$\begin{pmatrix} f_1 \\ \\ f_2 \end{pmatrix} (0.t_1 t_2 t_3 \ldots t_n) = A_{t_1} A_{t_2} A_{t_3} \ldots A_{t_n} \begin{pmatrix} 0 \\ \\ 0 \end{pmatrix} \tag{11}$$

where the operators A_0, A_1, A_2 and A_3 are defined by

$$\begin{aligned}
A_0 \begin{pmatrix} u \\ \\ v \end{pmatrix} &= \frac{1}{2} \begin{pmatrix} 0 & 1 \\ \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u \\ \\ v \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 \\ \\ 0 \end{pmatrix} \\
A_1 \begin{pmatrix} u \\ \\ v \end{pmatrix} &= \frac{1}{2} \begin{pmatrix} 1 & 0 \\ \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u \\ \\ v \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 \\ \\ 1 \end{pmatrix} \\
A_2 \begin{pmatrix} u \\ \\ v \end{pmatrix} &= \frac{1}{2} \begin{pmatrix} 0 & 1 \\ \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u \\ \\ v \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 \\ \\ 1 \end{pmatrix} \\
A_3 \begin{pmatrix} u \\ \\ v \end{pmatrix} &= \frac{1}{2} \begin{pmatrix} 0 & -1 \\ \\ -1 & 0 \end{pmatrix} \begin{pmatrix} u \\ \\ v \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 2 \\ \\ 1 \end{pmatrix}
\end{aligned}$$

The definitions of Hilbert curves of dimensions 3 and higher, and the definitions of *Peano* space filling curves, can be found in [16, 10].

Suppose that the segment $[0,1]$ is sampled at uniform intervals. With $T_{h_{2D}}$, if the number of sampling points is 4^m with m an integer, the points are mapped to a set of *grid points* that form a

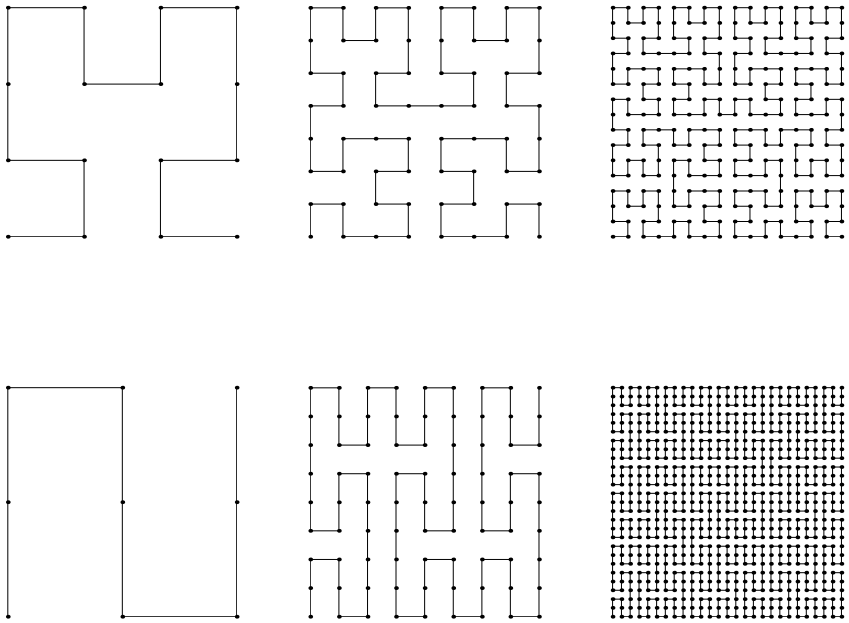

Figure 4: Top: Two dimensional Hilbert curves of orders 2,3 and 4. Bottom: Two dimensional Peano curves of orders 2,3 and 4. The grid points are marked.

uniform sampling of the unit square $[0,1]^2$ in a particular scanning order. We define the term *Hilbert curve of order m* to be the continuous curve obtained by connecting the resulting grid points in their scanning order. Let $T^m_{h_{2D}}$ denote the two dimensional Hilbert curve of order m. The Hilbert curve $T_{h_{2D}}$ can then be expressed as

$$T_{h_{2D}} = \lim_{m \to \infty} T^m_{h_{2D}} \ . \tag{12}$$

Two dimensional Hilbert and Peano curves of orders 2,3 and 4 are shown in Fig. 4.

Hilbert and Peano curves of a finite order have a *locality* property. Consider for example the 2-D 2nd order Hilbert curve shown in Fig. 5 (left) . Let R denote the distance between two adjacent grid points. Given any two points P, Q on the curve, let D_{PQ} denote the Euclidean distance between P and Q, and let d_{PQ} denote the distance between P and Q *measured along the curve*. The locality property in this case is that if $d_{PQ} > \sqrt{2}R$ then $D_{PQ} \geq R$. This means that if P and Q are sufficiently far away from one another along the curve, the Euclidean distance between them cannot be smaller

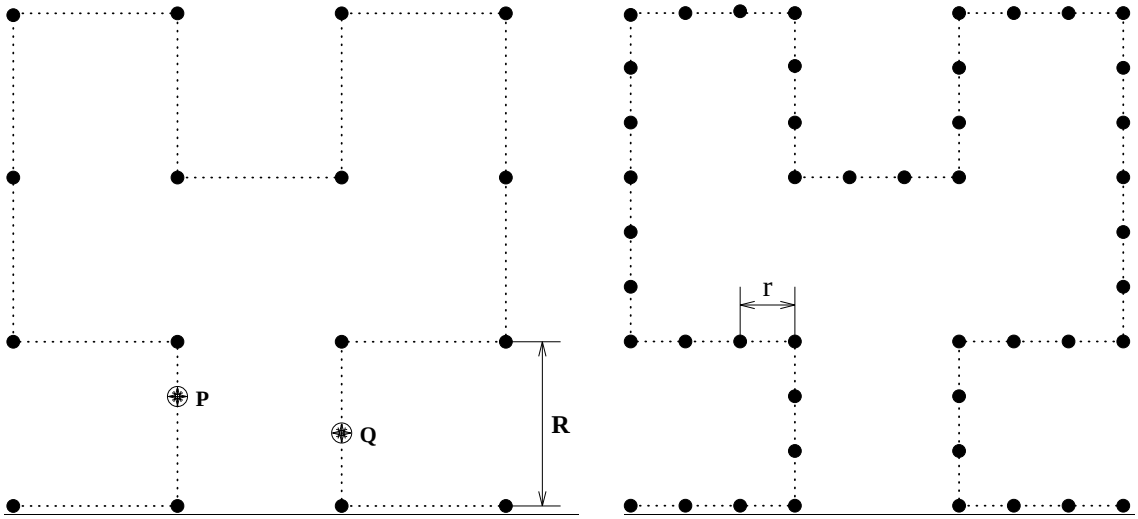

Figure 5: *Left:* The locality property in a two dimensional 2nd order Hilbert curve. R is the distance between adjacent grid points. The distance d_{PQ} measured along the curve is greater than $\sqrt{2}R$, hence the Euclidean distance D_{PQ} cannot be smaller than R. *Right:* The placement of 46 code points on a 2-D 2nd order Hilbert curve. Sixteen of the code points coincide with the grid points.

than R. For a theoretical discussion on locality in space filling curves see [9].

The number of grid points on a two dimensional Hilbert curve of order m is 4^m. Generally it is $c^{K \cdot m}$ where K is the dimension of the space and $c = 2$ for Hilbert curves, $c = 3$ for Peano curves. The number of grid points along each axis of $[0,1]^K$ is c^m, hence

$$R = \frac{1}{c^m - 1} \tag{13}$$

The length of a Hilbert (or Peano) curve of finite order m is thus

$$\Lambda = (c^{K \cdot m} - 1) \cdot R \approx c^{m(K-1)} \tag{14}$$

We suggest to sequentially place the L code points, equally spaced[1], along some space filling

[1]Slight improvement can be obtained by increasing the spacing near the corners of the curve.

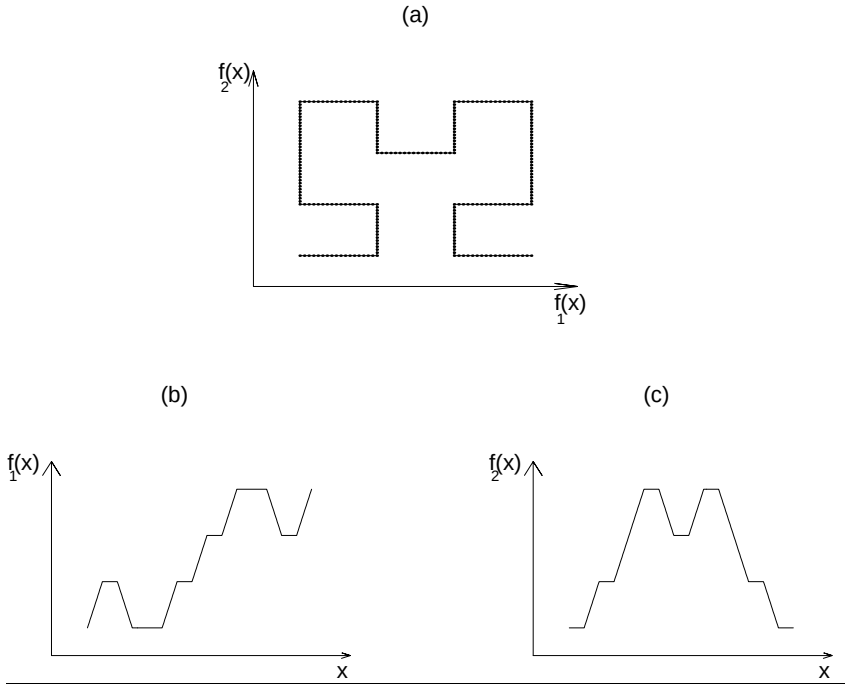

Figure 6: (a) A 2-D 2nd order Hilbert space filling curve. (b),(c) The illumination intensity profiles derived from that this word arrangement.

curve of finite order. This arrangement satisfies the main requirement, that any two code points that are distant from each other along the curve should be sufficiently distant in the code space. Spatially adjacent planes of light are thus represented by code words that are close to each other in the code space and distant planes of light correspond to distant code words. Figure 5 (right) shows the placement of 46 code points on a two dimensional 2nd order Hilbert curve. Figure 6 exemplifies (in the 2-D case) how the code vectors $\vec{f}(x)$ are derived from the curve. Code word assignment along a 3-D 2nd order Hilbert curve is shown in Fig. 7. The projection profiles derived from the curve and images of a scene under these illumination patterns are depicted in Fig. 8.

Let r denote the distance measured along the curve between two adjacent code points. If r is much larger than the average magnitude of the noise vectors in the system, the performance degradation due to noise will be minimal. Otherwise, the effects of noise will limit the effective resolution of the system. The number of light planes L that a structured light system projects is normally fixed by

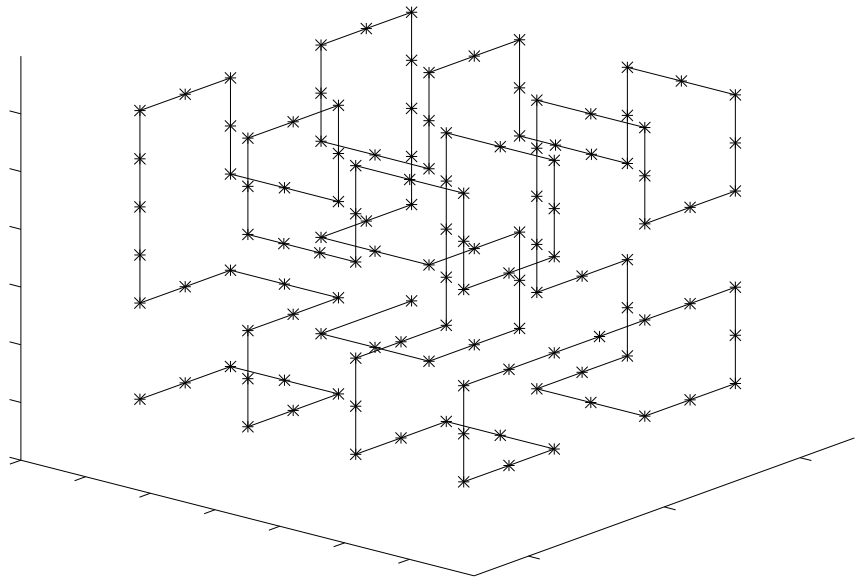

Figure 7: Code word assignment along a 3-D 2nd order Hilbert curve.

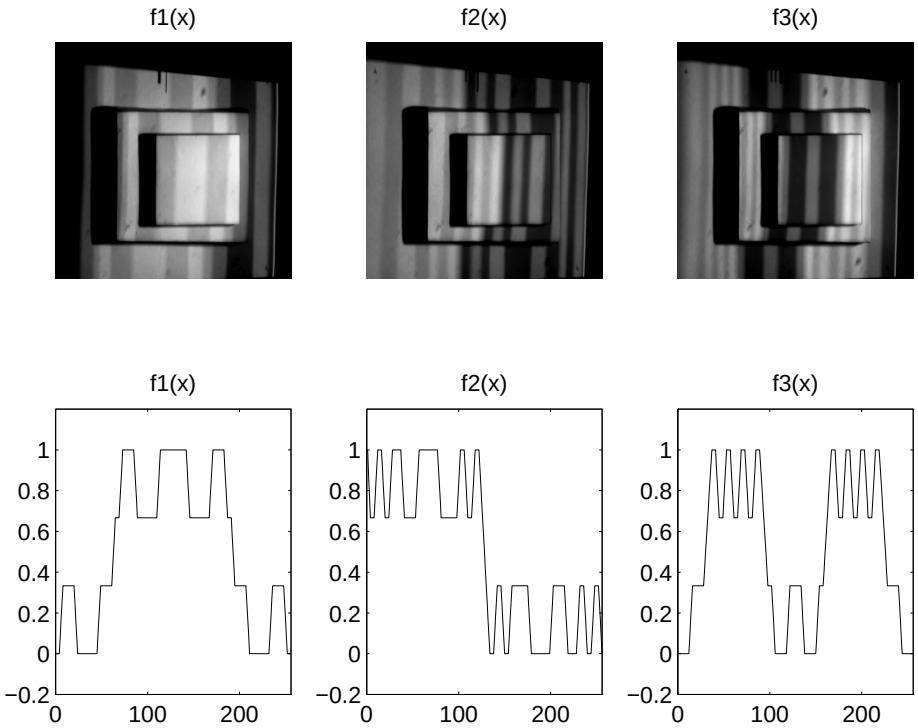

Figure 8: *Top:* A scene that is illuminated by projection patterns derived from the 3-D 2nd order Hilbert curve. *Bottom:* The corresponding illumination intensity profiles.

specifications and technology. The minimal useful value of r follows from the noise level. In order to accommodate L points that are r apart, the length Λ of the curve must be at least $L \cdot r$.

The length Λ of a Hilbert (or Peano) curve of order m can be increased by either increasing the order of the curve m (see Fig. 4), or by increasing the dimension of the code space K. Suppose that m is increased. Then R, the minimal distance between unrelated code words, decreases. At a certain point, the noise in the system will be able to skip the safety gaps frequently. Gross plane of light registration errors will then occur and the overall system performance will collapse. At this stage, the only way to usefully increase the length of the curve will be by increasing the dimension K, i.e, by increasing the number of illumination patterns.

In this framework the Gray coding scheme is simply a limiting case that should be used when the noise level is highest, the code words being placed as far as possible from each other at the vertices of a K dimensional unit hypercube.

4 Experiments

In order to exemplify the suggested approach and corroborate the theoretic results, we assembled a minimal structured light system that is roughly similar to the one described in [6]. It consists of a standard slide projector (Zett, Germany), a B/W CCD camera (Panasonic BL-200) and a PC with a frame grabber. 35mm slides were manufactured directly from computer files (8 bits/pixel) that contained the required illumination patterns. While this system is sufficient for the purposes of this research, it should not be compared, in terms of mechanical stability and speed, with current industrial structured light hardware.

The camera was positioned at a distance of about 130cm from the scene. The angle between the

optical axes of the camera and the projector was about 34^o. The geometric calibration procedure generally follows ref. [6]. All experiments were carried out under normal laboratory fluorescent light. This ambient light had little effect and no compensation was needed. Additional details on all technical aspects can be found in [10].

The standard deviation σ of the noise in intensity ratio measurements in the system was about 0.03 of full scale. It was fairly stable over a wide range of intensity ratio values. Hardware limitations dictated a maximal number $L = 256$ of illumination planes. Simulations indicated that robust discrimination of 256 illumination planes at the given noise level requires the dimension of the code space to be at least $K = 3$. This meant that at least three illumination patterns had to be projected, in addition to one or two uniform illumination references.

Consider a 3-D ($K = 3$) 2nd order ($m = 2$) Hilbert curve ($c = 2$). From Eq. 14, its length is $\Lambda = 21$. Placing 256 points at uniform spacing along the curve implies that the code space distance between adjacent light planes is $r \approx 0.082$, i.e., about 2.7σ. Disregarding the blending of adjacent light planes and assuming that the noise is Gaussian and statistically independent between pixels, this implies that 70% of the light planes will be perfectly registered, and that registration errors $|\hat{x} - x|$ larger than 1 are unlikely in a 256×256 image. Equation 13 indicates that the safety gap against gross errors is $R \approx 0.333$, i.e., more than 10σ.

What is the actual performance of these light plane labeling techniques? Figure 9 shows the standard deviation (in cm) of the depth in the reconstruction of a white plane, at a distance of about 135cm from the camera, using various labeling schemes. Each mark in Fig. 9 represents a specific method, and the horizontal axis shows K, the number of illumination patterns used in that scheme (excluding uniform reference illumination).

The nine marks connected by the dotted curve represent the standard Gray-code labeling scheme

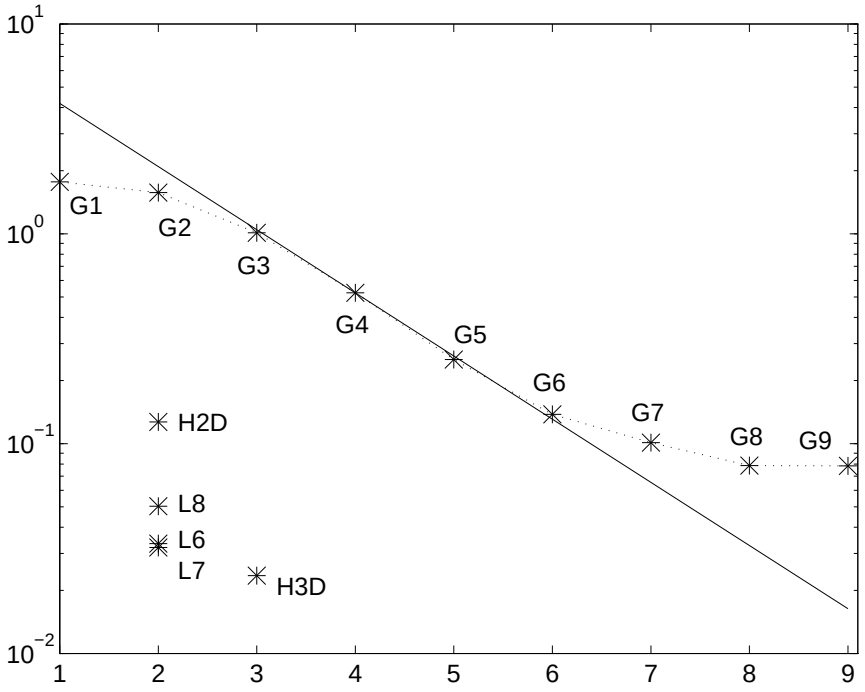

Figure 9: The standard deviation of the depth (cm) in the reconstruction of a white plane, at a distance of about 135cm from the camera, using various light-plane labeling methods. Each mark represents a specific labeling scheme. The horizontal axis shows K, the number of illumination patterns used in that scheme (excluding uniform reference illumination patterns). The nine marks on the dotted curve represent the standard Gray code technique with the indicated number of patterns. The labeling scheme derived from the 3-D Hilbert curve and used in all the reconstruction examples shown in the sequel (Figs. 10-12) is represented by the H3D mark. The H2D mark represents a two dimensional ($K = 2$) light plane labeling scheme derived from the 2-D 2nd order Hilbert curve. The L6, L7 and L8 marks represent 2-D schemes obtained with modified Peano curves [10] of orders 6,7 and 8 respectively.

with the indicated number of patterns. Ideally, since each additional Gray code bit doubles the number of distinct illumination planes, one expects these marks to lie on a straight line (shown solid) in a semi-log graph. Note that, for $K < 3$, the results are seemingly better than expected. This is an artifact created by clamping depth values that exceeded the workspace depth range to the depth limits of the workspace. Of greater interest is the saturation of the performance curve beyond $K > 6$. This can be explained by the limited optical bandwidth of the system. In certain regions in the scene, the projector is out of focus[2]. There the finer Gray code patterns are severely blurred and provide

[2]Generally, the camera might also be out of focus. In our system the illumination level made it possible to keep

little additional information. In a carefully designed industrial structured light system, the maximal useful value of K can be much larger than six. Yet, since the bandwidth of every optical system is limited, there will always be an upper physical bound on the improvement in depth measurement achievable by increasing the number of bits in the Gray code labeling scheme. A modified Gray code method that reduces the effective bandwidth of the projected patterns with respect to standard Gray code patterns has been suggested by Malz [13]. The illumination patterns that we suggest require a significantly smaller effective optical bandwidth than Gray code patterns.

The labeling scheme used in all the reconstruction examples shown in this paper (Figs. 10-13) is derived from the 3-D 2nd order Hilbert curve and represented by the H3D mark in Fig. 9. In our system it is superior to the Gray code labeling method not only at $K = 3$, but for all values of K. This means that in our experimental setup the performance of the H3D scheme (with 3 illumination patterns) cannot be matched by the Gray code scheme with *any* number of patterns. The H2D mark represents a two dimensional ($K = 2$) light plane labeling scheme derived from the 2-D 2nd order Hilbert curve. The L6, L7 and L8 marks represent 2-D schemes obtained with modified Peano curves [10] of order 6,7 and 8 respectively. They too are superior to the Gray code method in this case. The key to the superiority of the suggested methods with respect to the Gray code schemes is the better matching of the noise margins to the actual noise level in the system. When the noise level is extremely high, the Gray code scheme is obtained as a limiting case of the suggested approach.

In any real structured light system, the "light planes" have some thickness and, since a CCD cell is not a point, the "line of sight" is cone-like. Thus, in practice, the intersection of the two resembles a polyhedron rather than a point in space. This creates ambiguity and quantization errors that limit the resolution that can be achieved. The actual limiting values depend on the imaging geometry and

the camera aperture small and maintain sufficient depth of field. This is impractical in a standard commercial slide projector.

on the optics. When Gray code patterns are used in our system, the bound roughly corresponds to the solid line in Fig. 9. Note that the H3D, L6 and L7 schemes achieve error levels that are close to the error value on the solid line with $K = 8$ (ideal Gray code system with 256 illumination planes). This means that the measured performance of the H3D, L6 and L7 schemes, in this experiment, is close to the spatial ambiguity bound.

A significant error source in our experimental system is the misalignment of slides in the slide carrier within the slide projector. Since the manufacturer of a standard slide projector has no interest in maintaining tight mechanical tolerances in the slide carrier, there is often some misalignment between slides that leads to systematic errors. These errors can be avoided by using a mechanically fixed, computer controlled LCD spatial light modulator instead of changing slides.

A fundamental source of errors in all stereometric structured light systems is mutual illumination. If it cannot be assumed that the observed brightness of a point in the scene is proportional to the intensity of the light plane projected onto that point, identifying that plane of light becomes difficult. The significance of this problem is scene-dependent. In convex and flat parts of scenes it is often negligible, while at sharp concavities it might lead to gross errors. The reconstruction examples shown in Figs. 10-13 exemplify various aspects of system performance and limitations.

5 Conclusions

A theoretical framework for the optimal design of projection patterns for stereometric structured light systems has been presented. It draws on an analogy between the design of projection patterns and the design of signals for digital communication. Seemingly unrelated structured light methods, such as the Gray code scheme and intensity ratio techniques, are unified as special cases within the suggested framework.

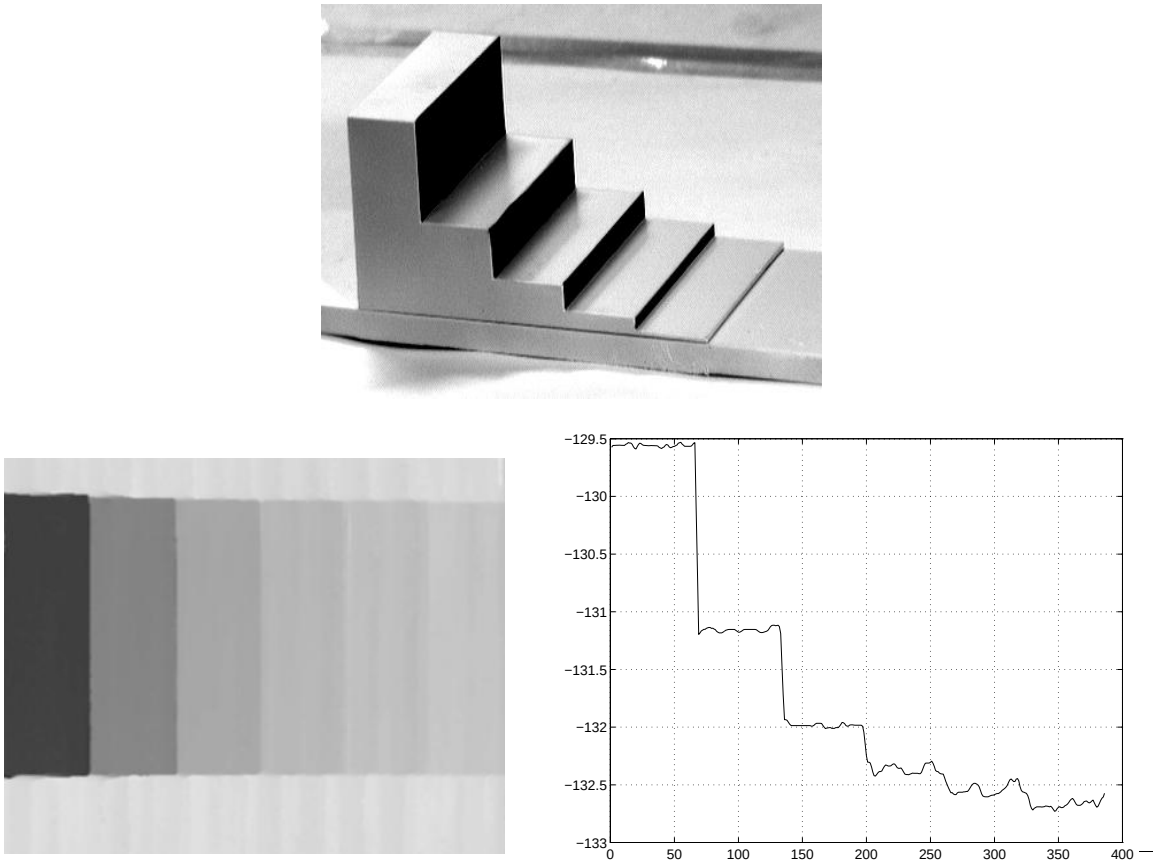

Figure 10: *Top:* An exponential staircase. The heights of the stairs are 16mm, 8mm, 4mm, 2mm and 1mm. *Bottom left:* The range image of the exponential staircase obtained at a distance of about 130cm from the camera. Dark points are closer to the camera. The total number of projection patterns used is five, including two uniform illumination references. *Bottom right:* A single row in the range map. Depth is in centimeters.

The design of K projection patterns for a structured light system with L distinct planes of light has been shown to be equivalent to the placement of L points in a K dimensional space subject to distance constraints imposed by the noise and errors. Since optimal design in the MSE sense seems to be an intractable multi-parameter global optimization problem, suboptimal solutions have been derived from families of K dimensional space-filling curves.

Our results indicate that performance similar to that of the Gray code technique can be reached with fewer projection patterns. This paves the way to faster acquisition and to a wider range of

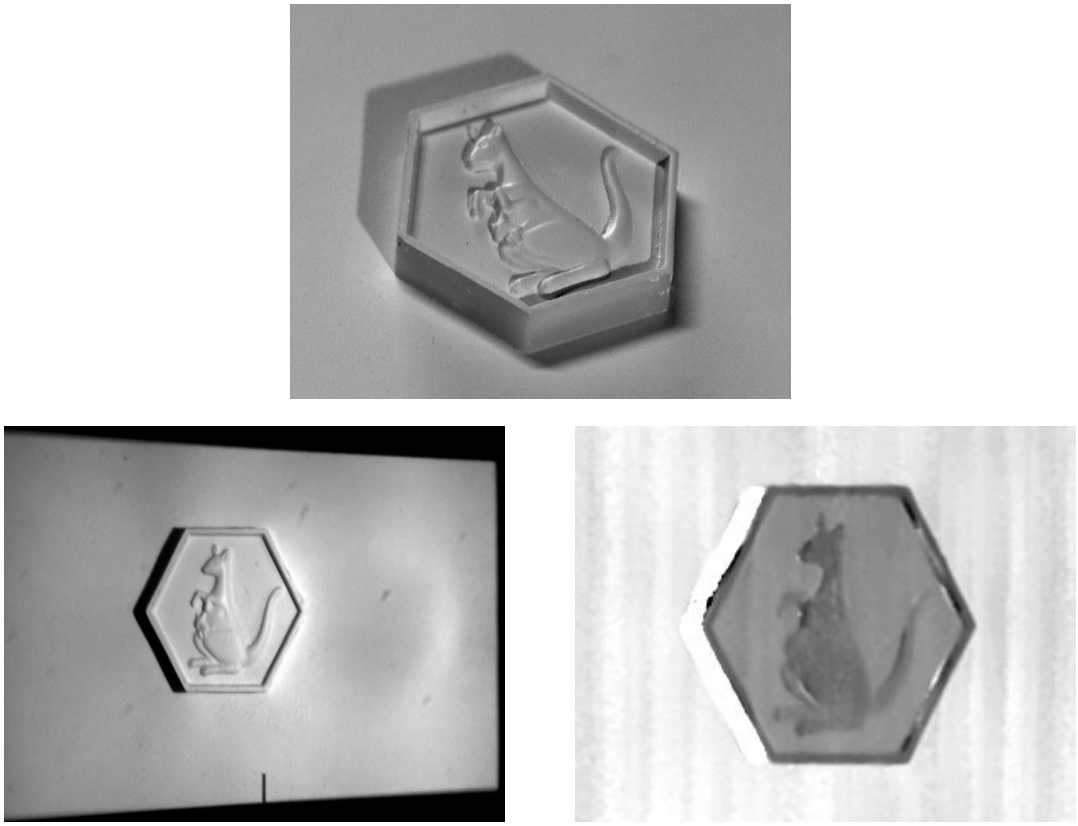

Figure 11: *Top:* A toy kangaroo coin. The thickness of the kangaroo relief is about 1.5mm. *Bottom left:* The image of the toy kangaroo coin taken by the system camera under uniform illumination. The scene is about 135cm away from the camera. Note the shadow region. *Bottom right:* The range image of the toy kangaroo coin. The details are clearly visible. The total number of projection patterns used is five, including two uniform illumination references.

applications. Further progress in this direction is reported in [7].

Acknowledgments

This research was supported by the Israel Ministry of Science and by the Ollendorff Center of the Electrical Engineering Department, Technion.

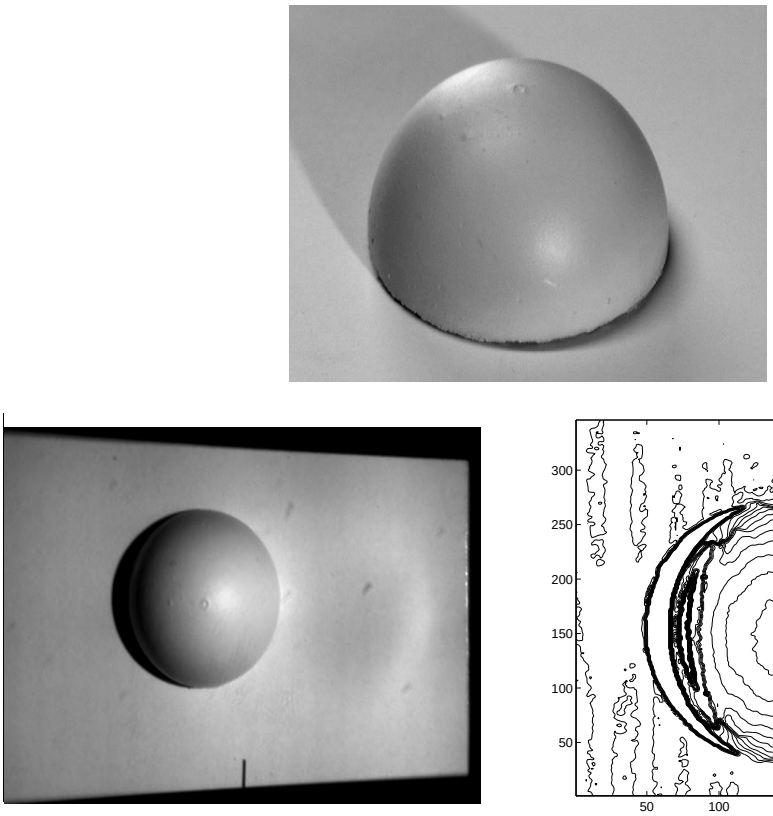

Figure 12: *Top:* A semi-sphere attached to a plane. The radius of the semi-sphere is 2.5cm. *Bottom left:* The image of the semi-sphere taken by the system camera, at a distance of about 135cm, under uniform illumination. Note the shadow region on one side of the semi-sphere, and the brightness of the edge on the other side, probably due to internal reflections (mutual illumination). *Bottom right:* Equal-height contour representation of the range image. Reconstruction is impossible in the shadow region due to lack of information, and distorted in the mutually illuminated region. Systematic errors, possibly due to misalignment of slides, can also be observed. The RMS error in sphere fitting is about 1mm.

References

[1] Agin G J and Binford T O, Computer Description of Curved Objects, in Proc. 3rd IJCAI (Stanford, 1973) 629-640.

[2] Åstrom A, Smart Image Sensors, Ph.D. Thesis (Linköping University, Sweden, 1993).

[3] Besl P, Active Optical Range Imaging Sensors, Machine Vision and Applications 1 (1988) 127-152.

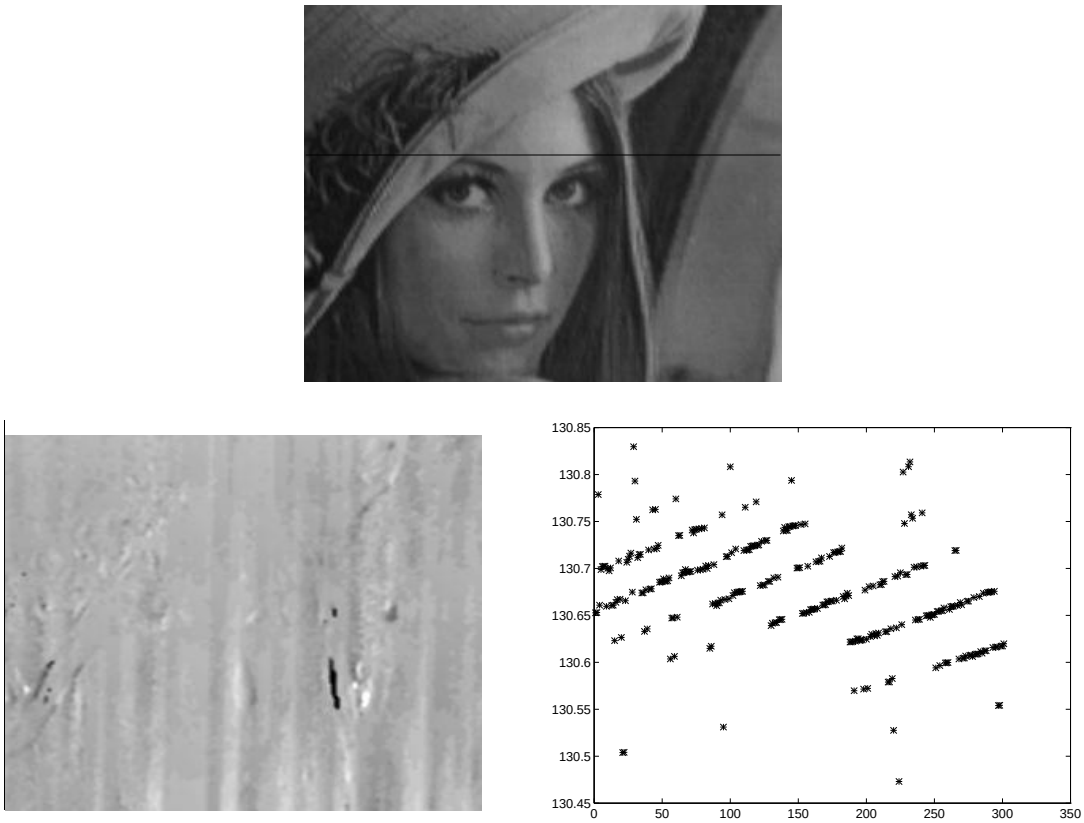

Figure 13: Reconstruction of a flat picture of Lena, positioned at a distance of about 130cm from the camera. *Top:* The image of the picture of Lena taken by the system camera under uniform illumination. One row is highlighted. *Bottom left:* A high resolution range image of the flat picture. White represents depth of 131cm, black 130cm. *Bottom right:* The depth reconstruction of the highlighted row. It can be seen that a single quantization step corresponds to about 0.5mm. The RMS depth error in this reconstruction is also about 0.5mm.

[4] Blake A, McCowen D, Lo H R and Lindsey P J, Trinocular Active Range Sensing, IEEE Trans. Pattern Anal. Mach. Intell. 15 (1993) 477-483.

[5] Boyer K L and Kak A C, Color-Coded Structured Light for Rapid Image Ranging, IEEE Trans. Pattern Anal. Mach. Intell. 9 (1987) 14-28.

[6] Carrihill B and Hummel R, Experiments with the Intensity Ratio Depth Sensor, Comput. Vis. Graph. and Image Process. 32 (1985) 337-358.

[7] Caspi D, Kiryati N and Shamir J, Adaptive Color Structured Light, Technical Report No. 1105 (Department of Electrical Engineering, Technion, Haifa, Israel, 1997).

[8] Chazan G and Kiryati N, Pyramidal Intensity Ratio Depth Sensor, Technical Report No. 121 (Center for Communication and Information Technologies, Department of Electrical Engineering, Technion, Haifa, Israel, 1995).

[9] Gotsman C and Lindenbaum M, On the Metric Properties of Discrete Space-Filling Curves, in Proc. 12th ICPR, III (Jerusalem, 1994) 98-102.

[10] Horn E, Optimal Structured Light Patterns, M.Sc. Thesis (Dept. of Electrical Engineering, Technion, Haifa, Israel, 1997).

[11] Kanade T, Gruss A and Carley L R, A Very Fast VLSI Rangefinder, in Proc. Int. Conf. on Robotics and Automation (1991) 1322-1329.

[12] Le Moigne J and Waxman A M, Projected Light Grids for Short Range Navigation of Autonomous Robots, in Proc. 7th ICPR (Montreal, 1984) 203-206.

[13] Malz R W, High Dynamic Codes, Self Calibration and Autonomous 3D-Sensor Orientation: Three Steps towards Fast Optical Reverse Engineering without Mechanical CMMs, in Proc. Optical Measurement Techniques III (Vienna, 1995) 194-202.

[14] Maruyama M and Abe S, Range Sensing by Projecting Multiple Slits with Random Cuts, IEEE Trans. Pattern Anal. Mach. Intell. 15 (1993) 647-651.

[15] Posdamer J L and Altschuler M D, Surface Measurement by Space Encoded Projected Beam systems, Comput. Graph. Image Process. 18 (1982) 1-17.

[16] Sagan H, Space Filling Curves (Springer, New York, 1994).

[17] Sato K and Inokuchi S, Three-Dimensional Surface Measurement by Space Encoding Range Imaging, J. of Robotic Systems 2 (1985) 27-39.

[18] Sato Y, Kitagawa H and Fujita H, Shape Measurement of Curved Objects Using Multiple Slit-Ray Projections, IEEE Trans. on Pattern Analysis and Machine Intelligence 4 (1982).

[19] Shirai Y and Suwa M, Recognition of Polyhedrons with a Range Finder, in Proc. 2nd IJCAI (London, 1971) 80-87.

[20] Smutný V, Analysis of Rainbow Range Finder Errors, in Hlavác V and Pajdla T eds., Proc. 1st Czech Pattern Recognition Workshop (1993) 59-66.

[21] Tajima J and Iwakawa M, 3-D Data Acquisition by Rainbow Range Finder, in Proc. 10th ICPR (Atlantic City, 1990) 309-313.

[22] Vuylsteke P and Oostrelick A, Range Image Acquisition with a Single Binary-Encoded Light Pattern, IEEE Trans. Pattern Anal. Mach. Intell. 12 (1990) 148-164.