

Physical Panoramic Pyramid and Noise Sensitivity in Pyramids

Weihong Yin* and Terrance E. Boulton†
Electrical Engineering and Computer Science Department
Lehigh University, Bethlehem, PA 18015

Abstract

Multi-resolution techniques have been used in a wide range of vision applications. Unfortunately, the costly operation of building a proper pyramid strongly reduces its value as a tool for reducing computational cost. A new approach, physical panoramic pyramid, is introduced in this paper. Physical panoramic pyramid measures multiple resolutions simultaneously resulting in multi-resolution panoramic images. No computation is needed to construct these image pyramids. We also analyze general noise sensitivity in image pyramids, including the interaction of the loss of resolution, random background noise and aliasing noise. The paper also discusses the issue of indexing between the neighboring layer, the viewpoint variation and the applications of the physical panoramic pyramid.

1 Introduction

There is a large body of research on multi-resolution and scale-space image processing and computer vision [1] [2] [4], and with the recent advances in wavelets the amount of research has redoubled to the point of multiple conferences on wavelets and applications per year, e.g. SPIE's [6]. Multi-resolution techniques, pyramid algorithms, have been widely used in vision applications such as segmentation, edge detection, motion estimation and tracking. Throughout the literature, three reasons dominate the justification for multi-resolution processing:

1. reducing computation via focus of attention and coarse-to-fine processing,
2. unknown scale or inherently multi-scale process such as edge detection or region segmentation,
3. its apparent relation to the human visual system.

A final advantage, the reduction of noise at the higher levels of the pyramid, may contribute to the success of multi-resolution algorithms. However, it is generally not explicitly stated as a motivating factor, and we are unaware of any formal studies on its impact.

While the pyramid algorithms have much to offer, they have had limited use in near-real-time tasks, because building a proper pyramid is a potentially costly operation requiring a prefiltering convolution before down sampling. For instance, presuming a 7×7 separable Gaussian convolution, we need 14 multiplies and 13 additions for each of $3 \times 640 \times 480$ pixels at $30fps$ for approximately $750MOPS$ to make the first layer of the pyramid, with $1/4$ of that for each additional layer, the total cost is about $1000MIPS$ just to form the pyramid. Although this can be done with today's processors, it is quite taxing and leaves few spare cycles for the actual processing. It becomes more significant when we consider HDTV or larger images, for which the large data rates demand "intelligent" processing. Building a good pyramid would require $12,000MIPS$ for HDTV and $70,000MIPS$ for $4k \times 4k$ video rate imagery.

To handle the computational burden, researchers have designed, built, and fielded, so-called pyramid machines[7]. These machines use multiple processors in parallel to produce a pyramid at video rate for 640×480 video. These pyramid machines also have the advantage of allowing parallel processing on the data of each level which is becoming important for algorithms such as image stabilization where an affine parameter is estimated for each patch at each level. However these specialized machines are costly and require considerable expertise to program and have significant Size Weight And Power(SWAP).

In [8] [9], S. Nayar revolutionized wide-field of view imaging by introducing an omni-directional sensor - a system that images a full hemisphere while allowing one to generate geometrically correct perspective images from the measured image. This paper extends the omni-directional sensor to physical panoramic pyramid using a set of parabolic mirrors. When using panoramic pyramid with a conventional camera/digitizer, we only pay to transfer the data which can be done via standard DMA, using only minimal CPU effort but still using system bandwidth and interfering with main memory access. With a CMOS camera or with cameras that support sub-frame mode access (e.g. CID cameras and many of the high resolution CCD) even the transfer of uninteresting data can be avoided. When the frame-grabber is on the other side of a slow bus or non-

*wey2@lehigh.edu

†tboulton@eecs.lehigh.edu. This work supported in part by ONR MURI contract N00014-95-1-0601.

DMA supporting interface (e.g. PCMCIA or ISA), this selective data access is equally important.

The paper is organized as follows. Section 2 describes the physical panoramic pyramid. Section 3 presents the noise sensitivity analysis in pyramids, including the interaction of the loss of resolution, random background noise and aliasing. Section 4 discusses the issue of indexing and viewpoint variation of the panoramic pyramid. Section 5 presents the applications of the panoramic pyramid and summarizes the paper.

2 Physical Panoramic Pyramid

In a para-camera[8] [9], a parabolic mirror is imaged by an orthographic lens to produce an omni-directional image. The combination of orthographic projection and the parabolic mirror provides a single viewpoint, at the focus of the parabolic surface. The image of the mirror, called the para-image, contains a hemi-spherical field-of-view, independent of the mirror size. The physical panoramic pyramid uses a set of parabolic mirrors stacked one on the top of the other. Figure 1 shows a three layer panoramic pyramid, where the mirrors were chosen so that the generated omni-directional images were 1/2 the resolution of the next finer resolution. Actually, mirrors can provide any resolution reduction desired. e.g. 4 to 1, 10 to 1 or even 6.4 to 1 (to reduce a $4k \times 3k$ image to normal 640×480 video). Figure 2 shows a panoramic pyramid image, the ratio of resolution between different image levels is 1:2:4. The “edges” of the mirrors only distort 1 pixel. To help understanding the scale, we note that the person is approximately 1 meter from the camera, the open door is 2 meters, the two 17" computer monitors(lower right and left) are 3 and 3.5 meter respectively. While the example in 1 shows a three layer panoramic pyramid with a 2:1 reduction rate and field-of-view of 360×45 degrees, we can choose to use only 2 mirrors, with a 4:1 reduction rate, which results in a 360×63.4 degree FOV.¹ The lowest resolution pyramid level with both is, however, the same. This 4:1 pyramid also places less demanding depth-of-field constraints on the imaging system.

For a panoramic pyramid using an NTSC camera, the maximum spatial resolution along the horizon is $\frac{240 \text{ pixels} * 2 * \pi}{360 \text{ degrees}} = 4.2 \frac{\text{pixels}}{\text{degrees}}$ (5.1 for PAL). Note the “spatial resolution” of the image is not uniform. While it may seem counter intuitive, the spatial resolution of the omni-directional images is *greatest* along the horizon, just where objects are most distant. If we zoom in to show only a quarter of the pyramid we reduce the FOV to 90x55 (or 90x80) but double the horizontal resolution to 8.4 pixels per degree. For comparison, a regular camera with

¹With mirrors viewing below the horizon we can extend the FOV further to 360x80 for a 4:1 mirror.

a 90x65 degree FOV would have maximum resolution of $\frac{640 \text{ pixels}}{90} \text{ degrees} = 7.1$, or about 15% less than the panoramic pyramid. The panoramic pyramid, however, has lower resolution in the vertical direction and its resolution decreases higher in the view.

While one could unwarp the panoramic images to produce multi-resolution perspective images in different directions, and then apply the algorithms in their “natural space”, the unwarping would add computation and introduce added errors. Similar issues arose in our surveillance work, where we have shown the speed advantages to be gained by properly adapting/developing algorithms to work in the raw omni-directional image space[12].

In this case of the panoramic pyramid there is an offset in the mirrors which produces a small viewpoint variation. For the panoramic pyramid in Figure 1, the viewpoints of the upper layers of the pyramid are offset by just under 9.8mm and 4.9mm from the layer below them. In two-layer pyramid with a 4:1 reduction, it is even smaller. As shown in section 4, the impact on the generated images is insignificant and can be ignored.

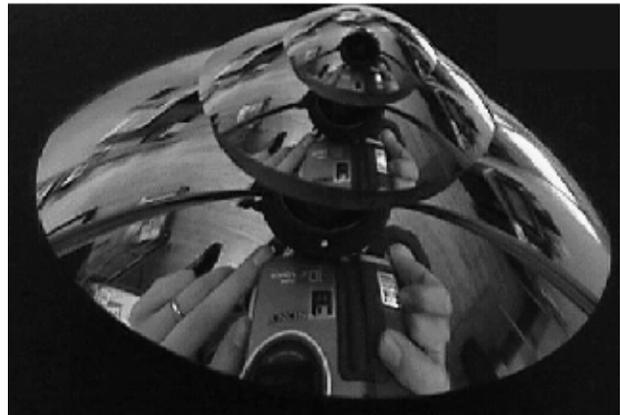


Figure 1. Three layer physical panoramic pyramid imaged from side to mirror stack

By using panoramic pyramid, algorithms are thus free to use coarse-to-fine focusing of attention in the truest sense, after processing the coarse level they transfer only the data of interest for the finer levels. In this way they significantly reduce not only the amount of the data processed but also the amount of data transferred. When looking at larger resolution imagers, this can be significant. For example uncompressed HDTV requires a transfer of *140MBytes* per second, more than the maximum bandwidth of the PCI bus.

Let us now consider the computational complexity of the different pyramid algorithms listed in Table 1. To construct traditional pyramid, we need multiplications, additions and loading operations. For example, building the first layer of the pyramid using 5×5 Gaussian kernel, re-



Figure 2. Multiresolution panoramic pyramid image, the ratio of resolution between different levels is 1:2:4. The original size is 480×480

quires $10N^2$ multiplications, $9N^2$ additions and N^2 loading operations, if the size of the image is $N \times N$. If we use ideal low-pass filter, we need $4Nn$ multiplications and $2Nn$ additions, where n represents $\log_2 N$, and another N^2 loading operations. If we use 2×2 block-averaging filter instead, only N^2 additions and N^2 loading operations are needed, but it can introduce significant aliasing artifacts (section 3). However, for physical panoramic pyramid, there is no computation cost and only the lowest resolution image needs to be loaded. For example, if we use a three-layer pyramid, the number of loading operation of the coarsest image is $N/4 \times N/4$. For problems where the computation on each level of the pyramid is simple, e.g.

segmentation or tracking, these savings can be significant. Take motion tracking for example, at each level we basically subtract a background image, which is computationally trivial. With panoramic pyramid, we may directly use the low-resolution para-image stream to detect the “blobs” and then use the only those parts of the high-resolution para-image needed to actually track detailed motions.

If the application requires standard perspective images, then the images from the highest resolution of the pyramid will have to be unwarped. This unwarping can introduce artifacts, but as was argued above, the resolution may actually be higher than a standard camera with similar FOV, so the warping introduced aliasing is not expected to be

Input Size	N^2	Level 1	Level 2
Output Size	$(N/2)^2$	512×512	256×256
Sample method		256×256	128×128
Gaussian $\sigma_g = 0.4$	$12N^2$	12×2^{18}	12×2^{16}
Gaussian $\sigma_g = 0.5$	$20N^2$	20×2^{18}	20×2^{16}
Gaussian $\sigma_g = 1.0$	$28N^2$	28×2^{18}	28×2^{16}
Ideal LP filter	$6Nn + N^2$	$27648 + 2^{18}$	$12288 + 2^{16}$
2×2 average	$2N^2$	2×2^{18}	2×2^{16}
Sample directly	N^2	2^{18}	2^{16}
Panoramic Pyramid	$(N/2)^2$	2^{16}	2^{14}

Table 1. The number of operations needed to obtain different pyramid levels. Note for panoramic pyramid, only the requested resolution image needs to be loaded

significant.

Another advantage of the panoramic pyramid is it is possible that the coarse resolution image can provide extra information that does not exist in the fine resolution image. For traditional pyramid construction, all the multi-resolution images are computed from the same original image, the coarse resolution image is only a simplification of the fine resolution image. However, the measurements of multiple resolutions images by panoramic pyramid are statistically independent. Combining this measurement can, at least in theory, reduce the camera noise.

Besides using panoramic pyramid, it is also possible to use normal CCD cameras combined with beam splitters to acquire multiple resolutions images, see Figure 3.

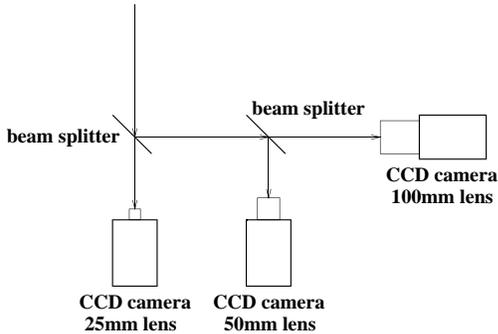


Figure 3. Use three beam splitters and CCD cameras to acquire multi-resolution images

3 Noise Sensitivity Analysis

3.1 Pyramid Simulation System

The most obvious advantage of pyramid representations is that they provide a possibility for reducing the computational cost of various image operations using coarse-to-fine strategy. To build the pyramid representation of an image, a smoothing process is applied followed by a subsampling operation. The properties of the smoothing filters

have been extensively studied by Burt[1] and Meer[3].

This filtering-sampling operation mainly has three effects: reducing resolution (or introducing blurring), reducing background random noise and introducing aliasing. If we also consider aliasing and non-ideal blurring as noise, there are three types of noise in each layer of pyramid images:

1. The noise introduced by the non-ideal blurring, N_b .
2. Aliasing noise, N_a , which is caused by subsampling.
3. The random background noise, N_n .

This paper studies the sensitivity of these different types of noise for different pyramid decomposition schemes.

To do this, a simulation model is constructed, which is illustrated in Figure 4. In this model, we have two signals as input, one is $F_0(n, m)$, the noise-free signal, the other is $I_0(n, m) = F_0(n, m) + N_0(n, m)$, where $N_0(n, m)$ is the random background noise which, for simplicity, is modeled as additive white Gaussian. The low-pass filters we studied are Gaussian filters, 2×2 block-averaging filter, ideal low-pass filter with band width π and directly subsampling. In digital cameras that provide reduced resolution in hardware, known as binning, block average down-sampling is used.

The ideal low-pass filter shown in the model is used to obtain non-aliasing signals. The upsampling process is performed in the frequency domain where the missing high frequency components are assigned as zero. In the model, i is the index of different pyramid layers. For the first level

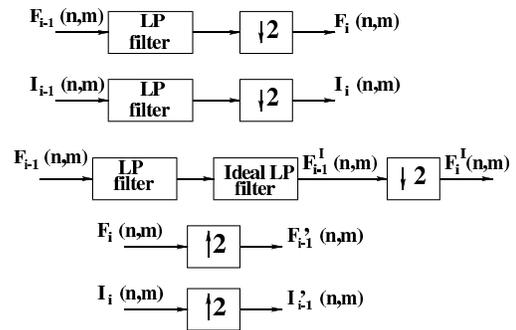


Figure 4. Simulation model for noise sensitivity of traditional pyramid

of the pyramid, the above three types of noise can be computed separately from this simulation model as:

$$\begin{aligned}
 N_{b1}(n, m) &= F_0^I(n, m) - F_0(n, m) \\
 N_{n1}(n, m) &= I_1(n, m) - F_1(n, m) \\
 N_{a1}(n, m) &= F_1(n, m) - F_1^I(n, m)
 \end{aligned} \tag{1}$$

where: $F_0^I(n, m)$ is the signal after low-pass filtering with its bandwidth cut to π ; $F_1(n, m)$ is the signal with aliasing

but no background noise; $I_1(n, m)$ has both aliasing and background noise; and $F_1^I(n, m)$ has neither aliasing nor background noise.

The corresponding $SNRs$ are defined as:

$$\begin{aligned} SN_{bR_1} &= 10\log_{10}\frac{\sigma_s^2}{\sigma_{b1}^2} \\ SN_nR_1 &= 10\log_{10}\frac{\sigma_s^2}{\sigma_{n1}^2} \\ SN_aR_1 &= 10\log_{10}\frac{\sigma_s^2}{\sigma_{a1}^2} \end{aligned} \quad (2)$$

where σ_s^2 is the variance of the original image F_0 , and σ_{b1}^2 , σ_{n1}^2 and σ_{a1}^2 are the variance of the N_{b1} , N_{n1} and N_{a1} respectively. Since the blurring noise and aliasing noise are not independent, we also consider their joint effect. As shown in Figure 4, $F_0'(n, m)$ is obtained by upsampling $F_1(n, m)$. It contains the joint artifact of blurring and aliasing. We denote the joint blurring and aliasing noise as N_{ba1} , which can be computed as

$$N_{ba1}(n, m) = F_0'(n, m) - F_0(n, m) \quad (3)$$

and the corresponding SN_{baR_1} is defined as:

$$SN_{baR_1} = 10\log_{10}\frac{\sigma_s^2}{\sigma_{ba1}^2} \quad (4)$$

Similarly, the overall noise N_1 at the first level is given by

$$N_1(n, m) = I_0'(n, m) - F_0(n, m) \quad (5)$$

where $I_0'(n, m)$ is obtained by upsampling $I_1(n, m)$, and contains all the three types of noise. Thus, the overall SNR_1 of the first layer of pyramid images is:

$$SNR_1 = 10\log_{10}\frac{\sigma_s^2}{\sigma_1^2} \quad (6)$$

For the upper levels ($i > 1$) of the pyramid, it is difficult to completely separate the blurring and aliasing artifacts. Thus, we only consider their joint effect with the random background noise and the overall noise. These can be obtained by:

$$\begin{aligned} N_{ba_i}(n, m) &= F_0'(n, m) - F_0(n, m) \\ N_{n_i}(n, m) &= I_i(n, m) - F_i(n, m) \\ N_i(n, m) &= I_0'(n, m) - F_0(n, m) \end{aligned} \quad (7)$$

where $F_0'(n, m)$ and $I_0'(n, m)$ are obtained by upsampling $F_i'(n, m)$ and $I_i'(n, m)$ through i levels. The corresponding $SNRs$ are defined as

$$\begin{aligned} SN_{baR_i} &= 10\log_{10}\frac{\sigma_s^2}{\sigma_{ba_i}^2} \\ SN_nR_i &= 10\log_{10}\frac{\sigma_s^2}{\sigma_{n_i}^2} \\ SNR_i &= 10\log_{10}\frac{\sigma_s^2}{\sigma_i^2} \end{aligned} \quad (8)$$

For physical panoramic pyramid, optics provide the “reduce” operation and introduce little aliasing noise. We are presuming, for now, that the proper optical design will result in a blur circle that is smaller than a pixel. The higher curvature and larger depth of field demands make this optical design more expensive than for the standard omnidirectional camera, but it is not considered too difficult. The current system does not satisfy the single pixel blur constraint, but before investing in development of the optics we undertook this simulation evaluation to insure the costs were warranted.

The background noise, which is modeling the random variations in the camera electronics, however occurs after the resolution reduction. We apply the method shown in Figure 5 to simulate the noise sensitivity of physical pyramids. An ideal low-pass filter or Gaussian low-pass filter with $\sigma_g = 1.0$ are used to approximate the optical “reduce” operator.² In keeping with the process model, we add per-pixel Gaussian noise *after* each blurring/subsampling operation. The computation of $SNRs$ is kept unchanged. Because we use ideal and Gaussian filters we can directly compare the impact of “post-pyramid” noise with the other artifacts.

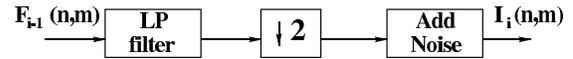


Figure 5. Simulation model for noise sensitivity of physical panoramic pyramid

3.2 Experiment Results

The evaluation used sixteen 8-bit 512×512 gray-level images, for of which are illustrated in Figure 6, see [13] for the others. For the background noise model each image was corrupted with additive random Gaussian noise with the standard deviations $\sigma_n = 0, \dots, 30$. The average of SNR over these 16 images is used to represent the noise sensitivity of traditional pyramids and the physical panoramic pyramid models. Table 3.2 shows the aver-



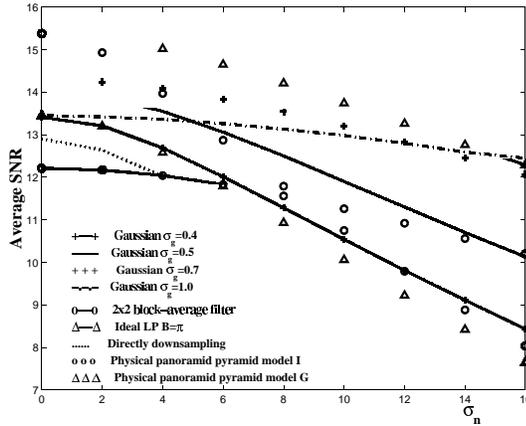
Figure 6. Four of the 16 test images

age SNR for the different pyramid algorithms when background Gaussian noise $\sigma_n = 2.0$. Based on our measurement, average standard deviation of background noise in the camera is around 2.0. The last two rows of the table are the $SNRs$ of the two physical panoramic pyramid models.

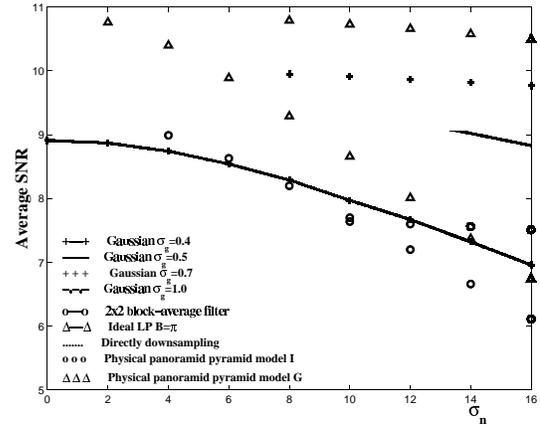
²A Besel or pill-box might be a more accurate model but would make comparison more difficult.

Sample method	level 1 (256 × 256)					level 2 (128 × 128)		
	SNR_1	$SN_{ba}R_1$	SN_bR_1	SN_nR_1	SN_aR_1	SNR_2	$SN_{ba}R_2$	SN_nR_2
Gaussian filter $\sigma_g = 0.4$	13.21	13.41	15.36	28.92	16.49	8.87	8.91	30.34
Gaussian filter $\sigma_g = 0.5$	13.90	14.04	15.22	31.35	18.50	9.75	9.77	35.12
Gaussian filter $\sigma_g = 0.7$	14.24	14.30	14.73	35.25	22.41	10.01	10.01	41.75
Gaussian filter $\sigma_g = 1.0$	13.42	13.45	13.68	38.47	26.57	9.31	9.31	45.40
2×2 block filter	12.17	12.21	12.63	33.51	20.55	7.72	7.73	39.55
Ideal LP filter	15.28	15.38	15.38	34.58	—	10.88	10.89	40.58
Sample directly	12.64	12.90	15.38	27.49	15.36	8.00	8.07	27.49
Physical pyramid model I	14.93	15.38	15.38	27.49	—	10.76	10.89	27.49
Physical pyramid model G	13.20	13.45	13.68	27.50	26.57	9.23	9.31	27.49

Table 2. Average SNR_s when $\sigma_n = 2.0$ (Average standard deviation of camera background noise in physical panoramic pyramid is around 2.0)



SNR in level 1



(b). Average SNR in level 2

Figure 7. Compare average overall SNR of different pyramid algorithms

One of them uses ideal low-pass filter, the other uses Gaussian low-pass filter with $\sigma_g = 1.0$. Figure 7 shows the changes of average overall SNR_i of different pyramid algorithms when σ_n of background Gaussian noise increases from 0.0 to 16.0.

From the results, we have the following general observations about noise effects in pyramids:

1. For the first level, where we could separate all noise components, it is clear that blurring dominated aliasing for all filters other than the ideal and direct downsampling. From the full data-set (not shown), we also find that non-ideal blurring is *the* dominant noise component when $\sigma_n \leq 8$. For level 2, for the traditional pyramid models the blur+aliasing noise dominates until $\sigma_n > 18$.
2. When $\sigma_n < 15$, ideal low-pass filter can provide the best performance among the different approaches that we studied here. This is due to the fact that blurring effect of the ideal low-pass filter is less than other filters and it does not introduce aliasing. But when $\sigma_n > 15$, the performance of Gaussian low-pass filter with $\sigma_g = 1.0$ is better than that of ideal low-pass filter because of its better background noise suppression

ability.

3. The background random noise is independent of blurring and aliasing. While, blurring noise and aliasing noise are highly correlated. In some images, for example, the $SN_{ba}R$ is even larger than the SN_bR .

We can also draw the following conclusions about the new physical pyramid models:

1. At level 1, the performances of two physical panoramic pyramid models are comparable to the pyramid algorithms using Gaussian low-pass filters and ideal low-pass filter, and is better than that of 2×2 block-averaging filter when σ_n of the background noise less than 6.0. (Recall our cameras have $\sigma_n = 2.0$.) When $6.0 < \sigma_n < 8.0$, the performance of the physical panoramic pyramid model using ideal low-pass filter is still better that of 2×2 block-averaging filter.
2. At Level 2, we see that for low and moderate noise levels, physical pyramids are better than 2×2 filtering, and for low noise they are better than Gaussian pyramids with small σ_g .
3. In all test cases, the new physical pyramid models are superior to direct downsampling, which is only pyra-

mid technique close in cost.

4 Error Analysis of Physical Panoramic Pyramid

As we mentioned before, the physical panoramic pyramid can directly measure multiple resolutions, the only computation is the user's algorithms being applied at the lower levels and then the indexing for the next finer level. This indexing has two components.

The first type of indexing is the generation of perspective views from the measured data. As in the case of the omni-directional images, this "unwarping" of the image can be reduced to a table lookup with optional interpolations such as nearest neighbor and linear interpolation [10].

The second indexing issue is relating the images at various levels of the pyramid to corresponding pixels at the next level. In traditional pyramids this can be done via a simple formula, for the panoramic pyramid the formulae are more involved but can be pre-computed from the sensor/mirror geometry. Furthermore, since the mirrors are stacked one on the top of the other, there is an issue of viewpoint variation. On the following derivations, we show that the impact of the viewpoint variation on the generated images and the computation of indexing between different levels are insignificant.

Initially, let us assume that the normal axes of two neighboring mirrors and their viewpoints are coincident, the ratio of the radius of two mirrors is $h_0/h_1 = R$, where h_0 is the radius of the big mirror, see Figure 8. A line in

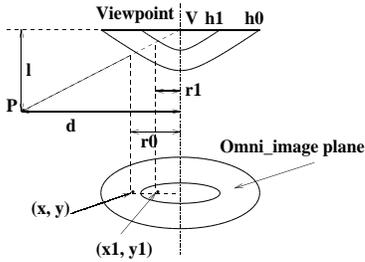


Figure 8. Two-layer panoramic pyramid, where the normal axes of two neighboring mirrors and their viewpoints are coincident

three-dimensional space will intersect with the paraboloid surface of the mirror at a distance ρ from its focus v :

$$\rho = \frac{h}{1 + \cos\theta}, \quad \cos\theta = \frac{l}{\sqrt{d^2 + l^2}} \quad (9)$$

and the projection of ρ on the para-image plane r can be described as:

$$r = \rho \sin\theta = \frac{hd}{l + \sqrt{d^2 + l^2}} \quad (10)$$

We observe:

$$\frac{r_0}{r_1} = \frac{h_0}{h_1} = R \quad (11)$$

so the relation between two projection points (x, y) and (x_1, y_1) on the two para-image planes is:

$$\frac{x - x_0}{x_1 - x_0} = \frac{y - y_0}{y_1 - y_0} = R \quad (12)$$

from equation(12), we obtain the initial indexing equations:

$$\begin{aligned} x_1 &= \frac{(R-1)x_0 + x}{R} \\ y_1 &= \frac{(R-1)y_0 + y}{R} \end{aligned} \quad (13)$$

Consider now the actual physical construction where the normal axes of two mirrors are coincident but there is a vertical distance $h_0/2$ (the height of the big mirror) between two viewpoints, see Figure 9. In this case, r_1

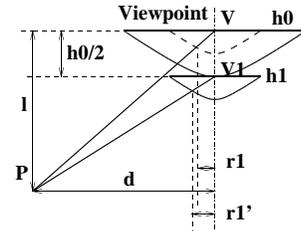


Figure 9. Two-layer panoramic pyramid, where the normal axes of two neighboring mirrors are still coincident and there is a vertical distance $h_0/2$ between two viewpoints of the mirrors

changes to r_1' :

$$r_1' = \frac{h_1 d}{(l - h_0/2) + \sqrt{d^2 + (l - h_0/2)^2}} \quad (14)$$

If we assume $l = 2m, d = 2m, h_1 = h_0/2$, and $h_0 = 12.7mm$. From equation (10) and (14), we obtain that $r_1 = 0.262mm$ and $r_1' = 0.269mm$. The difference between r_1 and r_1' is only $0.007mm$, which is around 0.14 pixel in the para-image. If $l = 1m, d = 1m$, The difference between r_1 and r_1' is $0.013mm$ and it is around 0.24 pixel in the para-image. So r_1' is approximately equals r_1 . Thus we can conclude the small variation of viewpoint in vertical direction can be ignored, except at very close range.

Finally we assume that there is horizontal shift between the two mirrors. In the para-image plane, there is translation between the centers of the projection of two mirrors:

$$\begin{aligned} \Delta x &= x'_0 - x_0 \\ \Delta y &= y'_0 - y_0 \end{aligned} \quad (15)$$

Handling this translation is straightforward. So the general indexing equations can be written as:

$$\begin{aligned} x_1 &= \frac{(R-1)x_0 + x}{R} + \Delta x \\ y_1 &= \frac{(R-1)y_0 + y}{R} + \Delta y \end{aligned} \quad (16)$$

Based on the above derivations, we conclude that the influence of the mis-indexing between two neighboring levels of the pyramid due to the small vertical variation of the viewpoint can be ignored. This type of mis-indexing is usually introduced by the height of the mirror. Meanwhile, the mis-indexing from the small horizontal variation of the viewpoint can be corrected by the shift of the viewpoint, which is easy to be measured.

The parameters needed in equation (16), x_0 , y_0 , R , Δx and Δy can be directly measured from the omni-image (Figure 2) and these variables are also required to generate the perspective views from the omni-image [10].

5 Discussion

Physical panoramic pyramid, which is inexpensive in computation, is an excellent alternative to traditional pyramid building algorithms. Multi-resolution omni-directional images can be obtained simultaneously using this approach. From the noise sensitivity analysis we see that physical panoramic pyramid are comparable to or better than the computationally constructed pyramids from low to moderate camera noise. We think the panoramic pyramid is a good alternative to the traditional multi-resolution approaches, especially for the real-time applications.

One of the ongoing research projects related to panoramic pyramid is multilevel color histogram representation of color images by peaks[11]. Where a two-level panoramic pyramid with a factor of 4 resolution reduction is used to get multi-resolution omni-images. In [11], it is shown that histogram peaks are more stable than general histogram bins where there are variation of scales. A room recognition system is also introduced which applies this indexing technique to omni-directional images of rooms.

The other research topic we are going to pursue is using panoramic pyramid on mobile robots. Our efforts are centered on algorithms for use in mobile-robot navigation. Because of the limited computational power of such systems we are starting with traditional NTSC/PAL based panoramic pyramid and developing hybrid algorithms for: location identification, flow-based obstacle avoidance, navigation, structure from motion and mosaicing/map building. At the same time we will also be testing/developing our optics and processing techniques for even larger format cameras, presuming that it will eventually become cost effective.

For motion tracking[12], We use low-resolution para-image stream to detect the objects and then using high-resolution para-image needed to actually track detailed motions.

While we have built a panoramic pyramid prototype, there are numerous research issues still to be addressed. The larger vertical extent of the stacked pyramids demands

a greater depth-of-field and more aggressive handling of field-curvature effects than is needed in standard omni-directional systems. As we move to higher-resolution, refined optical designs are needed to handle the smaller photo-site size and the larger total sensor size. Finally, even with the existing image systems the issues of flexible real-time access to the data will require considerable effort.

References

- [1] P.J. Burt, "Fast filter transforms for image processing," *Computer Graphics and Image Processing*, 16, pp. 20-51, 1981.
- [2] A. Rosenfield, editor, *Multiresolution Image Processing*, Springer-Verlag, New York, 1983.
- [3] P. Meer, S. Baugher, A. Rosenfield, "Optimal Image Pyramid Generating Kernels," *IEEE Trans. Pattern Anal. Machine Intel.*, Vol 9, 512-552, 1987.
- [4] T. Linderberg, *Scale-space theory in Computer Vision*, Kluwer Academic Publishers, 1994.
- [5] J.M. Jolion, A. Rosenfield, *A pyramid Framework for Early Vision*, Kluwer Academic Publishers, 1994.
- [6] SPIE, *Multiresolution Image Processing and Analysis*, V, 1995. Fifth in the series.
- [7] M. Hansen, P. Anandan, G. Van der Wal, K. Dana, P. Burt, "Real-time scene stabilization and mosaic construction," *Proc. of the IEEE WACV*, pp. 54-62, 1994.
- [8] S. K. Nayar, "Catadioptric Omnidirectional Video Camera," *Proc. of IEEE CVPR*, pp482-488, June 1997.
- [9] S. K. Nayar, S. Baker, "Catadioptric Image Formation," *Proc. of DARPA Image Understanding Workshop*, May 1997.
- [10] V. N. Peri, S.K. Nayar, "Generation of Perspective and Panoramic Video from Omnidirectional Video," *Proc. of DAPAR Image Understanding Workshop*, May 1997.
- [11] S. Sablak, T.Boult "Multilevel Color Histogram Representation of Color Image by Peaks for Omni-Camera," *Proc. of SIP'99*, Oct. 1999.
- [12] T.E.Boult, R.Michaels, X.Gao, P.Lewis, C.Power, W.Yin, A.Erkan, "Frame-Rate Omnidirectional Surveillance and Tracking of Camouflaged and Occluded Targets", *Second IEEE International Workshop on Visual Surveillance*, pp48-55, Fort Collins, Colorado,
- [13] W. Yin and T. Boult, "Panoramic Pyramids", Technical Report, Lehigh University, EECS Department. December 1998