Accurate lattice extraction in integral images

N. P. Sgouros, S. S. Athineos, M. S. Sangriotis, P. G. Papageorgas and N. G. Theofanous

Department of Informatics and Telecommunications, University of Athens
Panepistimiopolis, Athens, Greece, 15784
nsg@di.uoa.gr, sathin@di.uoa.gr, sagri@di.uoa.gr, ppapag@teipir.gr, optel@di.uoa.gr

Abstract: Integral imaging is one of the most promising techniques for delivering three-dimensional content. Most processing tasks usually require prior knowledge of the size and positions of the elemental images that comprise an integral image. In this paper we propose an automated method for calibrating the acquisition setup, by applying a preprocessing stage to an acquired integral image. The skew angle is extracted and the size and positions of the elemental images are accurately determined. For these purposes a method is developed to automatically identify an elemental image lattice that best matches the acquired integral image.

References and Links

1. G. Lippmann, "La Photographie Integrále," C.R. Acad. Sci. **146**, 446-451 (1908).
2. B.Javidi, F.Okano eds., Three-Dimensional Television, Video, and Display Technologies, (Springer-Verlag, Berlin, 2002), Chap. 4-6.
3. A.Stern and B.Javidi, "Three-Dimensional Image Sensing, Visualization, and Processing Using Integral Imaging," in *Proceedings of the IEEE* **94**, pp.591-607 (2006).
4. J. S. Jang, B. Javidi, "Formation of orthoscopic three dimensional real images in direct pickup one-step integral imaging," Opt. Eng. **42**, 1869-1870 (2003).
5. J. -H. Park, Y. Kim, J. Kim, S. -W. Min, and B. Lee, "Three-dimensional display scheme based on integral imaging with three-dimensional information processing," Opt. Express **12**, 6020-6032 (2004), http://www.opticsinfobase.org/abstract.cfm?URI=oe-12-24-6020.
6. M. Martinez-Corral, B. Javidi, R. Martínez-Cuenca, and G. Saavedra, "Formation of real, orthoscopic integral images by smart pixel mapping," Opt. Express **13**, 9175-9180 (2005), http://www.opticsinfobase.org/abstract.cfm?URI=oe-13-23-9175.
7. N.Sgouros, A.Andreou, M.Sangriotis, P.Papageorgas, D.Maroulis, N.Theofanous, "Compression of IP images for autostereoscopic 3D imaging applications," *in Proc. of IEEE International Symp. on Image and Signal and Image Processing* (IEEE, Rome, 2003) pp. 223-227.
8. S. Yeom, A. Stern, and B. Javidi, "Compression of 3D color integral images," Opt. Express **12**, 1632-1642 (2004), http://www.opticsinfobase.org/abstract.cfm?URI=oe-12-8-1632.
9. Y.Frauel and B.Javidi, "Digital three-dimensional image correlation by use of computer-reconstructed integral imaging," Appl. Opt. **41**, 5488-5496.
10. S.Hong and B.Javidi, "Improved resolution 3D object reconstruction using computational integral imaging with time multiplexing," Opt. Express **12**, 4579-4588 (2004), http://www.opticsinfobase.org/abstract.cfm?URI=oe-12-19-4579.
11. S.Yeom and B.Javidi, "Three-dimensional distortion-tolerant object recognition using integral imaging," Opt. Express **12**, 5795-5809 (2004), http://www.opticsinfobase.org/abstract.cfm?URI=oe-12-23-5795.
12. R.C.Gonzalez, R.E.Woods and S.L Eddins, *Digital image processing, using MATLAB* (Prentice Hall, NJ, 2004), Chap. 10.
13. T.H.Cormen, C.E.Leiserson, R.L.Rivest, *Introduction to algorithms*, (MIT Press, MA, 2000), Chap 16.
14. S. S. Athineos, N. P. Sgouros et.al, "Photorealistic Integral Photography using a Ray Traced Model of the Capturing Optics," J. Electron. Imaging (to be published).
15. R.C.Gonzalez and R.E.Woods, *Digital image processing, second ed.*(Prentice Hall, NJ, 2002), Chap. 3.
16. J. F. Canny, "A Computational Approach for Edge Detection," Trans. Pat. Anal. Mach. Intell. **8**, 679-698 (1986).
17. Y.Zheng, H.Li, D.Doermann, "A parallel-line detection algorithm based on HMM decoding," Trans. Pat. Anal. Mach. Intell. **27**, 777-792 (2005).

1. Introduction

The principles of integral imaging (InIm) were initially formulated by G. Lippman [1] back in 1908. The technique provides full color three-dimensional (3D) images, with adequate detail and depth levels and supports multiple simultaneous viewers [2]. Recent advances in InIm, limitations and performance optimizations are discussed in [3]. The basic principle for the acquisition setup consists of a charged coupled device (CCD) and a lens array [4,5], while a typical display setup is usually assembled of a liquid crystal display (LCD) and a suitable lens array, as depicted in Fig. 1(a) and 1(b) respectively. An InIm is realized as a collection of elemental images forming a two-dimensional (2D) lattice of small images. Once lens arrays with square elemental lenses are used in the acquisition setup the boundaries of the acquired elemental images create a 2D structure of perpendicular lines which we will call a lattice line structure (LLS), spanning the InIm.

Since both rotational and translational misalignments are present when the lens array is placed over the CCD, a calibration stage is needed to correct these aberrations. Even more, especially when using standard lens arrays and CCD sensors, the ratio between the lens array pitch and the CCD pitch takes non integer values. This fact causes a variability of the elemental image size that should be taken into account, when pseudoscopic to orthoscopic conversion [6] and InIm compression [7,8] is performed, as well as in all processing stages that usually require processing of each elemental image separately.
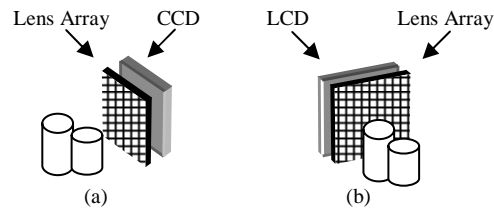

Fig. 1. Integral Imaging (a) acquisition and (b) display setup principles.

A number of different image processing techniques are used to locate corresponding points in neighbouring elemental images in order to address the problem of 3D object reconstruction from InIms' as reported in [9-11]. In this paper, we develop a scene independent method that automatically detects small rotational misalignments in InIm acquisition setups and measures the exact size and positioning of each elemental image without requiring prior knowledge on the system characteristics. In this way we can perform post acquisition processing in InIms or assist in the initial calibration of an InIm acquisition setup. Our work focuses on the accurate measurement of the skew angle of the LLS that reflects the rotational misalignment of the lens array used in the acquisition stage, in regard to the CCD sensor. It also manages to accurately detect the lattice lines positions, which usually deviate from the equidistant case due to the non integer values of the ratio of the lens array pitch to the CCD pitch. For this purpose we use non-linear filtering to enhance the boundaries of the elemental images and the Hough Transform (HT) [12] to produce an accurate estimate of the skew angle. A lattice matching method is developed, based on the techniques used to solve the longest common subsequence (LCS) problem [13], in order to match a detected sequence of lines positions to a theoretical lattice model. Our experimental data consist of a series of uncalibrated InIms acquired using one-step InIm [4] and computer generated InIms, using the method proposed in [14].

2. Description of the method

2.1 Skew angle detection stage

A color InIm is initially converted to a gray-scale image as shown in Fig. 2(a), prior to applying further processing. A 2D median filter [15] is used in order to compensate for noise and scratches that typically exist in physically acquired images. The Canny edge detector [16] is deployed next in order to detect strong lines in the InIm and produce the edge image depicted in Fig. 2(b). The image-wide lattice lines in the edge image are further enhanced using a one-dimensional (1D) median filter. On the contrary, shot noise and lines running to other directions are effectively attenuated due to their small size in regard to the median filter window. The filter is applied row-wise or column-wise and has the effect of attenuating everything except line segments that are almost horizontal or vertical respectively. In order to reduce the overall computational complexity of the algorithm, only one set of lattice lines, either horizontal or vertical, is considered in the skew detection process. As there is always a large number of elemental images in an InIm, there is also an adequate number of correctly detected lines running in a specific direction that is enough to produce an accurate estimation of the skew angle. The results of a vertical 1D median filtering are depicted in Fig. 2(c).

In the skew angle detection process we use the HT that is a robust and extensively used technique in image processing applications. The HT succeeds in identifying only image-wide lines, which are more likely to correspond to true lines of the LLS that define the boundaries of the elemental images. In our framework we used a sampling rate of 0.05^0 which provides a high level of accuracy, and assumed that the skew angle of the InIm remains under $\pm 10^0$ in order to reduce the computational complexity of the algorithm.

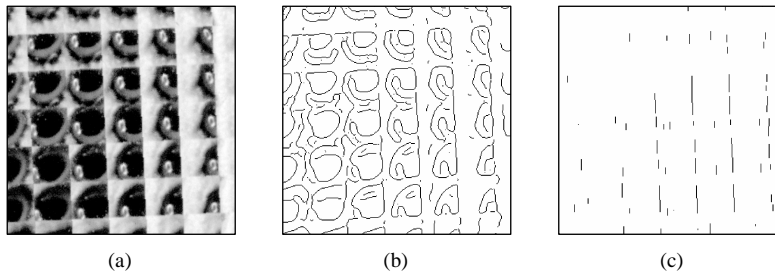(a)   (b)   (c)

Fig. 2. A portion of a physically acquired InIm. (a) Initial gray-scale image, (b) edge detection result, (c) edge image after vertical 1D median filtering.

The detection of the skew angle (θ_s) in the HT space is based on the fact that θ_s corresponds to a column in the HT accumulator array, **H**, having a large number of strong peaks, as shown in Fig. 3(a). On the contrary the rest of the columns in **H** are relatively homogenous having a small number of weak peaks. In order to determine θ_s we calculate the statistical variance of each column in **H,** as shown in Fig. 3(b).

2.2 LLS detection stage

In order to derive the lattice lines positions that form the LLS we apply the Canny edge detector and the 1D median filter on the deskewed version of the initial gray-scale InIm. Since the image is deskewed, the lattice lines are properly oriented in horizontal and vertical directions. In order to enhance these image-wide lines the 1D median filter is applied row-wise and column-wise creating the two images depicted in Fig. 4(b) and 4(c) respectively. Finally the horizontal and vertical projection profiles [17] are calculated for the images shown in Fig. 4(b) and 4(c) respectively. Peaks that are candidates for lattice lines in these two directions are identified using a global threshold at 20% of the maximum value of the corresponding profile. For each of the candidate peaks, a one dimensional version of the peak detection algorithm presented in [15] is used. The maximum number of expected peaks is

determined, based on the image dimensions and allowing for a minimum size of 10x10 pixels in each elemental image. A portion of the vertical projection profile and its thresholded version, corresponding to the filtered InIm from which a part is shown in Fig. 4(c), are depicted in Fig. 5(a) and 5(b).
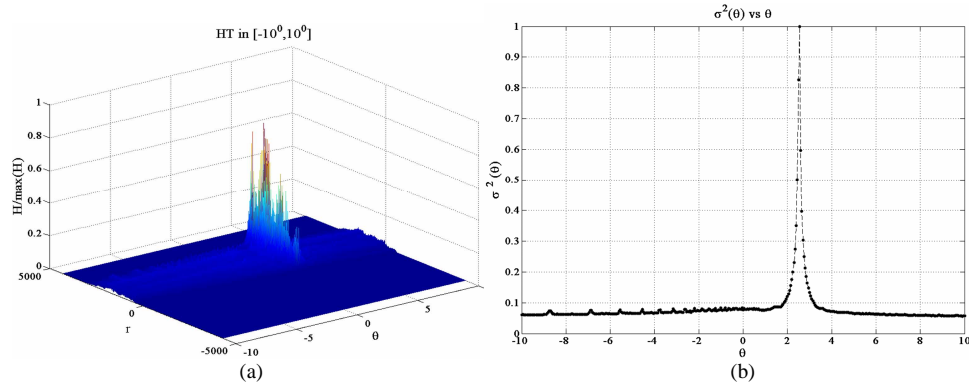

(a)

(b)

Fig. 3. (a) The column in the HT space corresponding to the skew angle θ_s has a large number of strong peaks, (b) the corresponding variance $\sigma^2(\theta)$ vs. θ.
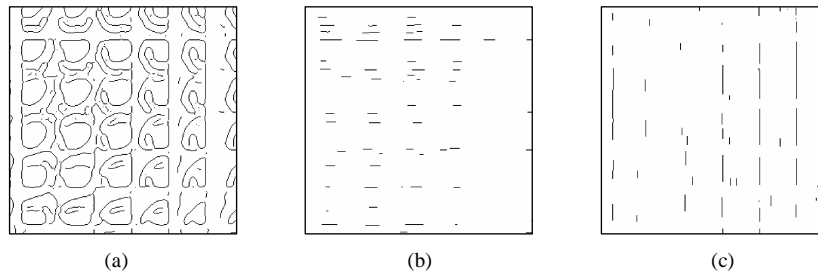

(a)   (b)   (c)

Fig. 4. (a) Deskewed part of an InIm, (b) Row-wise and (c) column-wise median filtering.
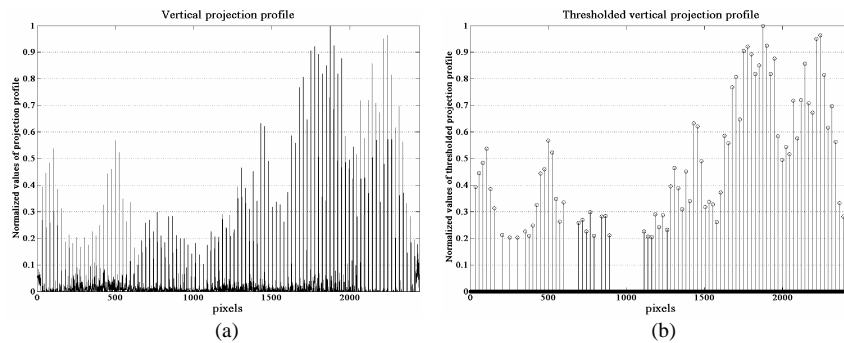

(a)   (b)

Fig. 5. Portion of (a) a vertical projection profile, (b) the corresponding thresholded projection profile.

In the previous stage we acquired a set of candidate locations for the lattice lines in the InIm. We denote these locations as two sequences $\{l_h\}=<l_{h1}, l_{h2},..., l_{hN}>$ and $\{l_v\}=<l_{v1}, l_{v2},..., l_{vM}>$ for the horizontal and vertical lattice lines respectively. A theoretical model that can be applied to predict the lattice lines locations is derived by using Eq. (1)

$$l_j = \left\lfloor \frac{p_0 + j \cdot p_l}{p_{ccd}} \right\rfloor \tag{1}$$

The initial offset p_0 and the lens pitch p_l can be expressed as $p_0 = k \cdot p_{ccd} + f_0$ and $p_l = n \cdot p_{ccd} + f_l$ respectively, where k is the integer number of CCD pixels in p_0, n the integer number of CCD pixels within each lens, $k < n$ and $0 \le f_0, f_l < p_{ccd}$. By substituting p_0, p_l in Eq. (1) we derive Eq. (2).

$$l_j = k + j \cdot n + \left\lfloor \frac{f_0 + j \cdot f_l}{p_{ccd}} \right\rfloor \tag{2}$$

By expressing f_l / p_{ccd} as a rational $q/r < 1$, with $\gcd(q,r) = 1$ and f_0 / p_{ccd} as $\lambda / r + e_0$, with λ taking integer values in $[0, r-1]$ and $0 \le e_0 < 1/r$, Eq. (2) finally resolves in Eq. (3).

$$l_j = k + j \cdot n + \left\lfloor \frac{\lambda}{r} + e_0 + j \frac{q}{r} \right\rfloor = k + j \cdot n + \left\lfloor \frac{\lambda}{r} + j \frac{q}{r} \right\rfloor \tag{3}$$

Using Eq. (3) we also derive the distance between consecutive lattice lines shown in Eq. (4).

$$d_j = l_{j+1} - l_j = n + \left\lfloor \frac{\lambda}{r} + (j+1) \frac{q}{r} \right\rfloor - \left\lfloor \frac{\lambda}{r} + j \frac{q}{r} \right\rfloor \tag{4}$$

It is easy to show that the expected interline distance calculated from the terms of $\{d\}$ is $m = E[d] = n + q/r$. Although Eq. (4) suggests that theoretically this histogram must have one or at most two populated adjacent bins, in Fig. 6 we observe a large number of occupied bins. This fact is due to the errors made in the line identification procedure. However in such a histogram there is always a single or two adjacent bins containing the largest part of the population. This situation is always valid for InIms that have an adequate number of correctly identified lines. From this histogram we can determine a confidence interval in which the mean value m is contained with a high probability. In order to reduce the error in the determination of the mean value m we form a joint histogram as the one shown in Fig. 6, by combining the values of the interline distances for the horizontal and vertical detected lines. In the case of only one strong bin with index d_p the confidence interval is chosen as $[d_p - \frac{1}{2}, d_p + \frac{1}{2}]$. Alternatively when two strong bins exist with indexes d_p and d_{p+1} this interval is chosen as $[d_p, d_p + 1]$. Using 11 equidistant points, including the intervals' boundaries, we provide a corresponding set of probable mean values, m_i, $i=1,2,\ldots11$. For each of these points we calculate the values n and q/r according to Eq. (5).

$$n_i = \lfloor m_i \rfloor \text{ and } \left(\frac{q}{r} \right)_i = m_i - \lfloor m_i \rfloor \tag{5}$$
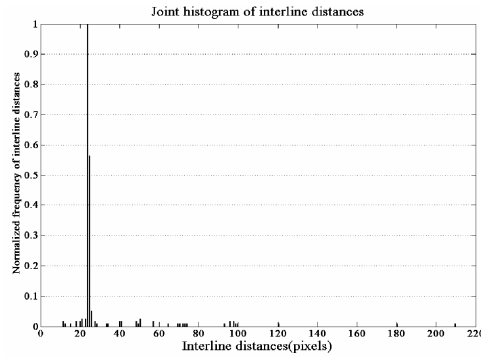
Fig. 6. Joint histogram of interline distances used to determine d_p

Using the values from Eq. (5) along with the corresponding value sets for λ and k it is easy to construct a number of different sequences $\{l\}$. These sequences are finally checked against $\{l_h\}$ and $\{l_v\}$ to determine a best match $\{l\}_{best,h}$, $\{l\}_{best,v}$ respectively. The best matches produce an LLS where the matching errors are minimized. In this work the mean value is determined with an accuracy of 0.1 pixels, whereas higher accuracy can be pursued at the expense of higher computational costs. In what follows we describe the matching process between the locations of the lines in each of these two sequences and the theoretical lattice lines model $\{l\}$ constructed for this purpose.

In order to obtain the best sequence $\{l\}_{best,h}$ amongst the set of all possible sequences $\{l\}$ that maximize a similarity function, each of the sequences $\{l\}$ is compared to $\{l_h\}$ and the longest common subsequence problem (LCS) is solved considering each pair of compared sequences as two strings of numbers. The longest run of consequent matches (LRCM) is calculated for each of these pairs created using $\{l_h\}$ and each of the sequences $\{l\}$. In our approach in order to find the LRCM we ignore single mismatches due to small gaps or spurious lines and penalize large consecutive mismatches. This process is repeated for determining $\{l\}_{best,v}$ that corresponds to the detected vertical lattice lines $\{l_v\}$. If $\{l\}_{best,h}$, $\{l\}_{best,v}$ correspond to the same value of m, these sequences are used separately to reconstruct the LLS. In the case that these sequences correspond to different values of m, an aggregating scheme is applied since square lenses are considered. The choice of the optimal match is based upon the sequence that has the greatest number of matches.

3. Experimental results

In order to evaluate the results of the method, we used real and computer simulated InIms to produce a number of different scenarios. Three different skew angles and lens array pitches were used in our simulation, in order to create InIms where each elemental image contained an average value of 14x14, 20.375x20.375 and 31.6x31.6 pixels. For the physical InIm acquisitions two different lens arrays were used resulting in InIms with an average elemental image pitch of 13.22x13.22 and 23.4x23.4 pixels. For the computer simulated InIms the skew angle and elemental image pitch size were determined beforehand, while for the physically acquired InIms, they were determined by visual inspection of the acquired image and the physical characteristics of the system. In total 15 samples were produced by varying the skew angle on these sets. The algorithm produced accurate results for over 80% of the above samples. Even in the cases where the method didn't find an exact estimate, it managed to generate a near optimal lattice that prohibited error propagation, and provided correct representation of the initial 3D scene. In Fig. 7(b) and 7(c) we show the results of an algorithmic pseudoscopy elimination procedure on the part of an InIm in Fig. 7(a) by using a constant pitch size or a variable pitch size determined with our method. In Fig. 7(b) and 7(c) we also provide magnified versions of the central segments of each image. The propagating

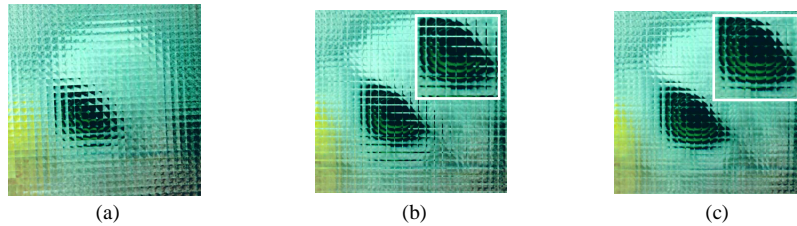errors when using a constant pitch size result in the badly restored orthoscopic InIm shown in Fig. 7(b).

| (a) | (b) | (c) |

Fig.7. (a) Uncalibrated portion of an InIm. Pseudoscopy elimination in the deskewed InIm, using (b) constant and (c) variable, pitch size.

4. Conclusions

In this paper we proposed an automated post acquisition calibration method, for use in InIm capturing setups. The robustness and flexibility of the method is based on the use of widely used image processing algorithms and a theoretical model to describe the lattice lines. Moreover the algorithm used in the matching process performs well even in cases where a large number of undetected or spurious lattice lines occur. The experimental results show that the method achieves good performance by successfully identifying the correct lattice in both physically acquired and synthetic InIms. Furthermore the results for the physically acquired InIms also indicate the effectiveness of the method in cases where noise, lens array imperfections or bad illumination conditions are present affecting the overall image quality.