Gamma/X-Ray Linear Pushbroom Stereo for 3D Cargo Inspection

Zhigang Zhu[*ab] , Yu-Chi Hu[bc] and Li Zhao[a]

[a]Department of Computer Science, The City College of New York, New York, NY 10031

[b]Department of Computer Science, The CUNY Graduate Center, New York, NY 10016

[c]Department of Medical Physics, Memorial Sloan-Kettering Cancer Center, NY 10021

[*] Contact Information:

Prof. Zhigang Zhu
Department of Computer Science
City College of New York /CUNY
Convent Avenue and 138th Street, New York, NY 10031
Tel: (212) 650 – 8799   Fax: (212) 650 - 6248
Email: zhu@cs.ccny.cuny.edu
URL: http://www-cs.engr.ccny.cuny.edu/~zhu/

ABSTRACT

For evaluating the contents of trucks, containers, cargo, and passenger vehicles by a non-intrusive gamma-ray or X-ray imaging system to determine the possible presence of contraband, three-dimensional (3D) measurements could provide more information than just 2D measurements. In this paper, a linear pushbroom scanning model is built for such a commonly used gamma-ray or x-ray cargo inspection system. Three-dimensional (3D) measurements of the objects inside a cargo can be obtained by effectively constructing a pushbroom stereo system using two such scanning systems with different scanning angles. A simple but robust calibration method is proposed to find the important parameters of the linear pushbroom sensors. Then, a fast stereo matching algorithm is developed to obtain 3D measurements of the objects under inspection. This algorithm is fully automatic based on free-form deformable registration. An interactive user interface is designed for 3D visualization of the objects of interest. Using the interactive tool, the automatic algorithm is also compared with a very simple semi-automatic algorithm based on point correlation. Experimental results of sensor calibration, stereo matching, 3D measurements and visualization of a 3D cargo container and the objects inside, are presented.

Keywords: Pushbroom imaging, automatic 3D measurements, stereo matching, cargo inspection, homeland security

1. INTRODUCTION

With the ongoing development of international trade, cargo/vehicle inspection becomes more and more important. Quite a few X-ray or gamma(γ)-ray inspection systems have been put into practical uses (Hardin 2002; Hardin 2004; Hitachi 2004, Orphan et al 2002). In this paper, a non-intrusive gamma-ray imaging system (Orphan et al 2002) will be used as an example to describe our research work. This system produces gamma-ray radiographic images, and has been used for the evaluation of the contents of trucks, containers, cargo, and passenger vehicles to determine the possible presence of many types of contraband. In the past, however, cargo inspection systems have only had two-dimensional capabilities, and human operators made most of the

visual measurements. But if we could build an accurate geometry model for the gamma-ray imaging system, which turns out to be a linear pushbroom scanning sensor as modeled by Gupta and Hartley (1997), three-dimensional (3D) measurements of objects inside a cargo container can be obtained when two such scanning systems with different scanning angles are used to construct a *linear pushbroom stereo system*. The 3D measurements add more value to today's cargo inspection techniques, as indicated in some recent online reports (Hardin 2002; Hardin 2004; Hitachi 2004).

Pushbroom images (or mosaics, when generated from video sequences) with parallel-perspective projections are very suitable for such surveillance and/or security applications where the motion of the sensor has a dominant translational direction. Examples include satellite pushbroom imaging (Gupta and Hartley 1997), airborne video surveillance (Zhu et al 2001, Zhu et al 2004), 3D reconstruction for image-based rendering (Chai and Shum 2000), road scene representations (Zheng and Tsuji 1992; Zhu and Hanson 2004), under-vehicle inspection (Dickson et al 2002; Koschan et al 2004), and 3D measurements of industrial parts by an X-ray scanning system (Gupta et al 1994; Noble et al 1995). A pushbroom image/mosaic is called a *parallel-perspective* image when there is a parallel projection in the direction of the sensor's translational motion and perspective projection in the direction perpendicular to that motion. A pair of pushbroom stereo images/mosaics can be used for both 3D viewing and 3D reconstruction when they are obtained from two different oblique viewing angles. An advantageous feature of the pushbroom stereo is that depth resolution is independent of depth (Chai and Shum 2000; Zhu et al 2001; Zhu, et al 2004). Therefore, better depth resolution could be achieved than with conventional perspective stereo, or the recently developed multi-perspective stereo with circular projection (Peleg et al 2001; Shum and Szeliski, 1999; Klette et al 2001; Wei & Klette, 2006), given the same image resolution. Rotating line cameras project 3D scenes on a cylindrical surface, where multiple rotating sensor lines allow to compose panoramic stereo images. We note that multi-perspective stereo with circular projection that is based on wide-baseline line cameras can achieve quite accurate depth resolution for far-range airborne scenes (Klette, et al, 2001; Huang). However, in such a configuration, depth resolution is still proportional to the square of depth; therefore depth accuracy varies significantly for the cargo inspection case with large depth variations. Furthermore, the circular motion that is required in such a setting is not the best form for scanning long cargo containers.

In this paper, issues on 3D measurements using a linear pushbroom stereo system are studied for gamma-ray or X-ray cargo inspection. The closest work to ours is the x-ray metrology for industrial quality assurance (Gupta et al 1994; Noble et al 1995). However, to our knowledge, our paper presents the first piece of work in using linear pushbroom stereo for 3D gamma-ray or X-ray inspection of large cargo containers, with fast and fully-automated 3D measurements. This paper uses the gamma-ray scanning images provided by the Science Applications International Corporation (SAIC) (Orphan et al 2002). Even though this paper talks specifically about gamma-ray, the algorithms developed in this paper can be used for pushbroom images acquired by X-ray or other scanning approaches as well.

In our previous work (Zhu et al 2005), we have developed the pushbroom stereo model and an effective calibration method for 3D cargo inspection. A simple, interactive correlation-based stereo matching procedure was also presented to validate the feasibility of 3D measurements for cargo inspection. In our more recent paper (Zhu and Hu 2007), we presented a new automated stereo matching method modified from a free-form deformable registration approach (Lu et al 2004), which is more suitable for gamma-ray images. Fast implementation and image enhancement are also incorporated into stereo matching. This paper is a combined and extended version of the above two conference presentations. In addition, we have also added new materials to compare the two stereo matching approaches - the interactive correlation-based and the fully automated ones, and to compare and integrate multiple pairs of pushbroom stereo images.

This paper is organized as follows. Section 2 introduces the geometry of the pushbroom scanning sensor model. In Section 3, the geometry of the pushbroom stereo system is developed. Then in Section 4, a calibration method is proposed to find the important parameters of the sensor model. In Section 5, a fast and automated stereo matching algorithm is developed considering both the radiographic and geometric properties of gamma-ray stereo images. To facilitate stereo matching, an adaptive window min-max method is used to enhance the contrasts and boundaries of the challenging gamma-ray images. For fast implementation, a multi-resolution approach is also applied. In Section 6, an interactive user interface is designed to visualize the 3D measurements of objects of interest. In this section, comparison is also made between the automated algorithm and a simple, correlation-based interactive algorithm using the designed

user interface. Multi-view comparison and integration is also discussed here. Finally, in Section 7, we conclude our work and discuss a few future directions in both research and applications.

2. γ-RAY LINEAR PUSHBROOM SENSOR MODEL

The system diagram of the gamma-ray cargo inspection system (Orphan et al 2002) is shown at the bottom-left corner of Figure 1. A 1D detector array of 256 NaI-PMT probes counts the gamma-ray photons passing through the vehicle/cargo under inspection from a gamma-ray point source. Either the vehicle/cargo or the gamma-ray system (the source and the detector) moves in a straight line in order to obtain a 2D scanning of gamma-ray images.
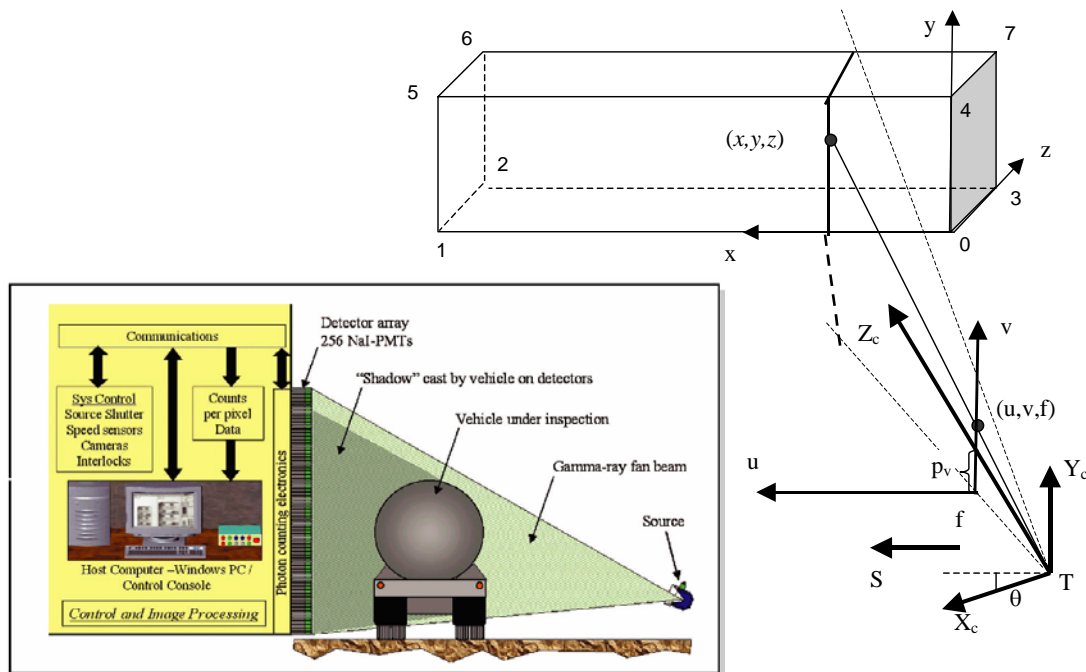

Figure 1. Linear pushbroom sensor model. A gamma-ray cargo inspection system (Courtesy SAIC, San Diego, CA, USA) that can be modeled by linear pushbroom geometry, is show in the bottom-left corner of the figure.

The geometry of the system is shown in Figure 1 (right). The 1D detector array geometry can be modeled by the well-known perspective projection camera $X_cY_cZ_c$ with the optical center at the location of gamma-ray source, and the 1D detector array in the vertical direction v and at a distance f along the optical axis Z_c (i.e., f is the focal length in pixels). Note that in the figure the

image is drawn between the objects and the optical center for easy illustration. The scanning begins when the optical center of the sensor is at location $T = (T_x, T_y, T_z)$ in the world coordinate system o-xyz. The angle between the optical axis (TZ_c) of the sensor and the oz axis of the world coordinate is θ. We assume that there are no tilt and roll angles between the two coordinate systems.
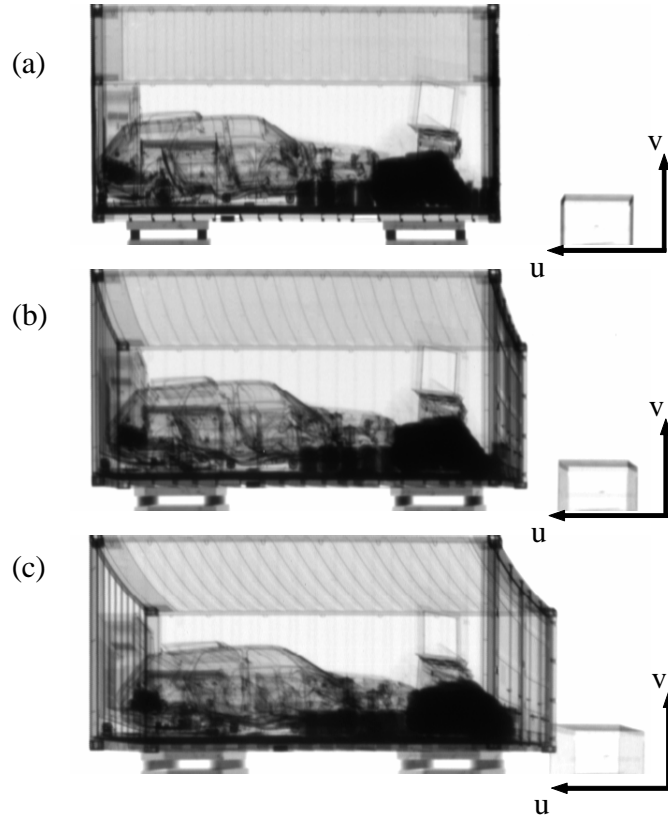

Figure 2.   Real gamma-ray images with three different scanning angles (a) zero (b) ten and (c) twenty degrees (Courtesy SAIC, San Diego, CA, USA).   Each image has a size of 621x256 pixels, i.e., 621 scans of the 256-pixel linear (column) images. The stereo visual displacements, particularly the back surface of the cargo container, are obvious by comparing the three images.

Assume that the sensor moves at a constant speed S (feet per vertical scan line, i.e., ft/pixel) in the direction of the x-axis, then the velocity vector represented in the camera coordinate system $X_cY_cZ_c$ is $V = (V_x, V_y, V_z) = (S\cos\theta, 0, S\sin\theta)$. The center of the linear image in the v direction is defined by a vertical offset p_v. Putting all of these into the linear pushbroom projection equation formulated by Gupta and Hartley (1997), we have the relationship between a 3D point (x,y,z) in the world coordinate system and its image point (u,v) as

$$
\begin{pmatrix} u \\ wv \\ w \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & f & p_v \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \dfrac{1}{Vx} & 0 & 0 \\ -\dfrac{Vy}{Vx} & 1 & 0 \\ -\dfrac{Vz}{Vx} & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix} \left(\begin{pmatrix} x \\ y \\ z \end{pmatrix} - \begin{pmatrix} Tx \\ Ty \\ Tz \end{pmatrix} \right)
$$

This linear pushbroom equation leads to the following two simpler equations:

$$
u = \frac{x - Tx - (z - Tz)\tan\theta}{S} \tag{1}
$$

and

$$
v = f\cos\theta \frac{y - Ty}{z - Tz} + p_v \tag{2}
$$

Note that the pushbroom scanning system has parallel projection in the u direction (Eq. (1)), but has perspective geometry in the v direction (Eq. (2)). Figure 2 shows three real gamma-ray images, with three different scanning angles – zero, ten and twenty degrees, respectively. Each image has a size of 621x256 pixels, i.e., 621 scans of the 256-pixel linear column images.

3. γ-RAY LINEAR PUSHBROOM STEREO

A dual-scanning system is a *linear pushbroom stereovision system*. It can be constructed with two approaches: two simultaneous linear pushbroom scanning sensors with different scanning angles, or a single scanning sensor to scan the same cargo twice with two different scanning angles. For a 3D point *(x,y,z)*, its image correspondences in the stereo pair can be represented by

$$
\begin{aligned}
u_k &= \frac{x - T_{xk} - (z - T_{zk})\tan\theta_k}{S_k} \\
v_k &= f_k \cos\theta_k \frac{y - T_{yk}}{z - T_{zk}} + p_{vk}
\end{aligned} \qquad (k=1, 2) \tag{3}
$$

Therefore the depth of the point can be recovered as

$$
z = \frac{d - d_0}{\tan\theta_1 - \tan\theta_2} \tag{4}
$$

where

$$d = S_2 u_2 - S_1 u_1 \tag{5}$$

is the *visual displacement* (measured in feet) of the point (*x, y, z*) measured in the pair of stereo images, and

$$d_0 = (T_{x1} - T_{z1} \tan \theta_1) - (T_{x2} - T_{z2} \tan \theta_2)$$

is the fixed offset between the two images. In Figure 2, the visual displacements, particularly the back surface of the cargo container, are obvious by comparing the three images. Note that in Eq. (4), the depth z is obtained by only using the u coordinate of a point pair in the stereo images (Eq. (5)). Further, the depth of any point is proportional to its visual displacement in the stereo pair. Therefore the depth resolution is independent of depth.

After the depth is obtained via pushbroom stereo, the x and y coordinates of the point can be calculated from one of the two images, for example

$$x = u_1 S_1 + z \tan \theta_1 + T_{x1} - T_{z1} \tan \theta_1$$
$$y = \frac{(v_1 - p_{v1})(z - T_{z1})}{f_1 \cos \theta_1} + T_{y1} \tag{6}$$

4. SENSOR CALIBRATION

In order to use two scanning systems to calculate 3D information, we need to calibrate each scanning system first. For each scanning setting, the following parameters are required for 3D estimation: the focal length f, the image center p_v, the scanning angle θ, the scanning speed S, and the initial sensor location (T_x, T_y, T_z). In order to fulfill this task, we need to know a set of 3D points $\{(x_i, y_i, z_i), i=1, 2,\ldots, N\}$ and their corresponding image points $\{(u_i, v_i), i=1, 2,\ldots, N\}$. Our calibration method only needs to know the dimension of the container, which is

length(x)*height(y)*depth(z)=20*8*8 (ft^3).

in the example of Figure 2. The 3D coordinates of the eight corners of the container can be easily calculated automatically using the shape model of the container. Then we locate the 8 vertices of

the rectangular container (refer to Figure 1) in each gamma-ray image by manually picking up the 8 corresponding image points.

An interesting property of the linear pushbroom sensor is that Eq. (1) and Eq. (2) can work independently. Therefore, in calibrating the sensor, we first obtain the "parallel projection parameters" in Eq. (1) and then the "perspective projection parameters" in Eq. (2). Eq. (1) can be turned into a linear equation with three unknowns, i.e., S, $tan\theta$ and T_x-$T_z tan\theta$:

$$u_i S + z_i \tan \theta + (T_x - T_z \tan \theta) = x_i \qquad (7)$$

Given more than three pairs of points ($i=1, 2,..., N$ where $N \geq 3$), we can solve the linear system to find the three unknowns by using the least square method. Similarly, Eq. (2) leads to a linear equation with five unknowns, i.e. f, fT_y, p_v, $p_v T_z$ and T_z:

$$(y_i \cos \theta) f - \cos \theta (fT_y) + z_i p_v - (p_v T_z) + v_i T_z = v_i z_i \qquad (8)$$

With the known θ and given more than five pairs of points ($i=1, 2,..., N$ where $N \geq 5$), we can solve the linear equation system. Note that from Eq. (7) we can only find the values of the speed S and the angle θ and a combined parameter T_x-$T_z tan\theta$. Nevertheless, this is sufficient for obtaining the depths of points using Eq. (4).

Table 1 shows the results of the "parallel parameters" for all the three settings corresponding to images a, b, and c in Figure 2. The remaining parameters, including Tx, can be obtained after solving Eq. (8), which is needed to calculate the x and y coordinates of 3D points by using Eq. (6). Table 2 shows the "perspective parameters" and the T_x values for all the three settings.

Table 3 shows the 3D measurements using the eight *image* point pairs used for calibration between the ten-degree and twenty-degree images. The purpose is to examine the accuracy of the pushbroom stereo modeling and calibration results. The numbers of the points listed in Table 3 are labeled in Figure 1 for comparison. For the container with a dimension of 20x8x8 ft^3, the average errors in depth z, length x and height y are 0.064 ft, 0.033 ft and 0.178 ft respectively, indicating that the linear pushbroom modeling and calibration method is sufficiently accurate for 3D measurements. Note that the accuracy of the estimation in Table 3 reflects the errors in both the linear pushbroom sensor modeling and the sensor calibration, including image point

localization. In theory, the depth error δz introduced by image localization error δu can be estimated as the first derivative of z with respect to u using Eqs. (3) and (4), that is

$$\delta z = \frac{S}{\tan\theta_1 - \tan\theta_2}\,\delta d \tag{9}$$

To simplify the formula (Eq. (9)), we assume that the two scans have the same speed (i.e. $S_1 = S_2 = S$), which are almost true for our example in Figure 2 (see Table 1). In this example, one-pixel image localization error introduces an error of 0.254 ft in depth estimation, using the parameters listed in Table 1.

Table 1. Parallel projection parameters

Img	S (ft/pixel)	Tanθ	θ (degrees)	T_x-$T_z$tanθ
A	0.04584	0.00143	0.0821	-7.398
B	0.04566	0.16552	9.3986	-7.283
c	0.04561	0.34493	19.031	-7.309

Table 2. Perspective projection parameters

Img	F (pxl)	T_y (ft)	p_v(pxl)	$p_v T_z$	T_z (ft)	T_x(ft)
A	427.78	-0.41558	21.148	-177.78	-14.815	-7.419
B	441.24	-0.42881	17.787	-191.78	-15.141	-9.789
C	456.18	-0.41037	19.250	-198.03	-15.000	-12.48

Table 3. 3D measurements of the test points

No	x	y	z	dx	dy	dz
0	-0.033	-0.179	-0.063	-0.033	-0.179	-0.063
1	20.033	-0.177	0.063	0.033	-0.177	0.063
2	19.967	-0.152	7.936	-0.033	-0.152	0.064
3	0.033	-0.204	8.064	0.033	-0.204	0.064
4	-0.033	7.787	-0.063	-0.033	-0.213	-0.063
5	20.033	7.856	0.063	0.033	-0.144	0.063
6	19.967	7.799	7.936	-0.033	-0.201	0.064
7	0.033	7.844	8.064	0.033	-0.156	0.064

We have the following notes about the calibrations results:

(1) The parallel parameters estimation: the estimated speeds for scanning the three images are almost the same (S = 0.0456-0.0458), and the angles obtained are very close to the parameters provided by SAIC, i.e., 0, 10 and 20 degrees.

(2) The perspective parameters estimation: the three sets of parameters, including the focal lengths, the image centers, and the camera initial locations are consistent with each other.

(3) The parallel parameters are more accurate than the perspective ones due to fewer parameters in calibration and no inter-dependency among unknowns in the former, whereas three of the five unknowns in Eq. (8) are not independent, which includes parameters that relate the y and v coordinates, thus creating larger errors in the estimations of the y coordinates than the x coordinates (Table 3). In solving Eq. (8) using SVD, we found that one of the four singular values of the matrix $\mathbf{A^T A}$ is almost zero, where $\mathbf{A}$ is the coefficient matrix of the linear system of Eq. (8). Therefore, the pseudo inverse of matrix $\mathbf{A^T A}$ was used.

(4) The poor accuracy and a system bias in the y direction could also be caused by the inaccurate input image point locations, particularly in the v direction. It is hard to localize the actual corresponding points of the eight corners of the containers in the gamma ray images. In the images shown in Fig. 2, the points corresponding to the corners #0 and #1 are mixed with the background and therefore hard to identify, whereas the points #4 and #5 are right at the first row of the images therefore their more accurate positions could possibly be out of the view.

5. AUTOMATED STEREO MATCHING

The gamma-ray images in cargo inspection are similar to those X-ray images generated by a medical imaging system. Therefore registration techniques using in medical images (Lu et al 2004; Xu 2000) could be employed for our application. We adapted a free-form deformation registration method (Xu 2000) for our gamma-ray stereo matching. There are several advantages of this method. First, it is automatic: it is a pixel-based registration method so it does not require any feature extraction and can be done fully automatically. Second, it is fast: it is capable of finding local deformation, i.e., the displacement field for each pixel, and hence a global minimization can be conducted efficiently via calculus of variations (Keener 1998; Xu 2000).

Finally this method, originally designed for x-ray image registration, is appropriate for the γ-ray images having the similar properties.

The free-form deformable registration problem is described as finding the displacement field of a pair of images that minimizes an energy function. The energy function is composed of not only the similarity of intensities of two images, but also the smoothness of the displacement field. Let A($\mathbf{u}$) be the reference image and B($\mathbf{u}$) be the target image, the displacement is defined as a mapping from A to B:

$$\mathbf{d_u} : A \to B \tag{10}$$

So that a point $\mathbf{u} = (u_1, u_2) = (u, v)$ in the reference image moves to $\mathbf{u} + \mathbf{d_u}(\mathbf{u})$ in the target image. The energy function is defined as

$$\varepsilon(\mathbf{d_u}) = \int_{\mathbf{u} \in \Omega^2} \left[R^2(\mathbf{u}, \mathbf{d_u}) + \lambda \sum_{i=1}^{2} \sum_{j=1}^{2} \left(e_{ij} \right)^2 \right] d\mathbf{u} \tag{11}$$

Here $R(\mathbf{u}, \mathbf{d_u}) = B(\mathbf{u} + \mathbf{d_u}) - A(\mathbf{u})$ is the residual between the two images, λ is a constant weight, and $e_{ij} \equiv \dfrac{\partial d_{u_i}}{\partial u_j}$ $(i = 1,2; j = 1,2)$ is the partial derivative of the displacement vector which describes how smooth the movement is around a pixel. The method tries to minimize the difference of the intensity while maintains the smoothness of the displacement fields at the same time. We can see that when R^2 is small, the energy is dominated by summation of the squares of the partial derivatives of the displacement vector, yielding a smoothly varying field.

Since the energy function is a function of variables $\mathbf{u}$, $\mathbf{d_u}$ and $\dfrac{\partial \mathbf{d_u}}{\partial \mathbf{u}}$, by using the calculus of variations (Keener 1998; Xu 2000), the displacement field can be found by solving the following Euler-Lagrange equation

$$\lambda \nabla^2 \mathbf{d_u} - R(\mathbf{u}, \mathbf{d_u}) \frac{\partial R(\mathbf{u}, \mathbf{d_u})}{\partial \mathbf{d_u}} = 0 \tag{12}$$

where $\dfrac{\partial R(\mathbf{u},\mathbf{d_u})}{\partial \mathbf{d_u}}$ is the gradient vector field of the deformed image. Then a finite difference scheme is used to solve this nonlinear elliptic partial differential equation, and Newton iteration is applied to update the displacement iteratively, as

$$d_{uk}^{new} = d_{uk}^{old} + \frac{L_k^{old}}{\lambda + (g_k^{old})^2}, k = 1,2 \qquad (13)$$

Here $L_k = \lambda \nabla^2 d_{uk} - [B(\mathbf{u}+\mathbf{d_u}) - A(\mathbf{u})]g_k(\mathbf{u}+\mathbf{d_u})$ and g_k is the gradient in the k's dimension $(k = 1,2)$ of the image B at $\mathbf{u}+\mathbf{d_u}$.

A multi-resolution technique is applied in our implementation. It can be viewed as a systematic way for structuring local information into global information in order to deal with the gamma-ray images that lack salient visual features. Using a multi-resolution technique also overcomes the limitation of a one-resolution approach in handling large displacements using the free-form deformable registration method. In the multi-resolution registration, large displacements can be found with computational efficiency at coarser level, even with lower accuracy, which is later refined with the finer resolution calculation. Four layers of image resolutions are used in our implementation: 621x256, 310x128, 155x64 and 77x32. From coarsest level to the finest resolution, the numbers of iterations to find the displacements at each level are 128, 64, 32 and 16, respectively. After registration at a lower level is done, the displacements at this level then are up-sampled and scaled to the next higher resolution level and to be used as the initial displacements for that level.

We also utilize the epipolar geometry constraint of the pushbroom stereo system in stereo matching. In a stereo imaging system, the displacement vector of a point should be on the point's epipolar line. The epipolar lines of a stereo pair from the linear pushdown stereo scanning system are approximately horizontal scanlines (i.e., $v_1 \approx v_2$ in Eq. (3)). Thus in the multi-resolution registration process, we limit the displacement vector to be one-dimensional in the horizontal direction (i.e., along the x-axis) at all levels except the finest resolution level. At the finest resolution level, we allow an offset of +2/-2 pixels (depending on the accuracy of the sensor modeling) in the vertical direction (i.e., along the y-axis) to account for geometric errors in sensor setup and sensor calibration.
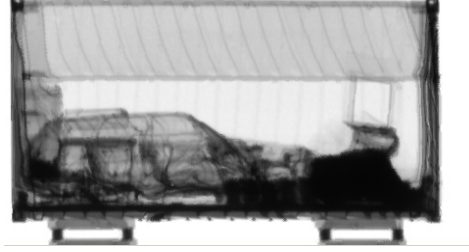
Figure 3. Warping the target image (Figure 2b) to the corresponding reference image (Figure 2a) using the estimated displacement vector fields.
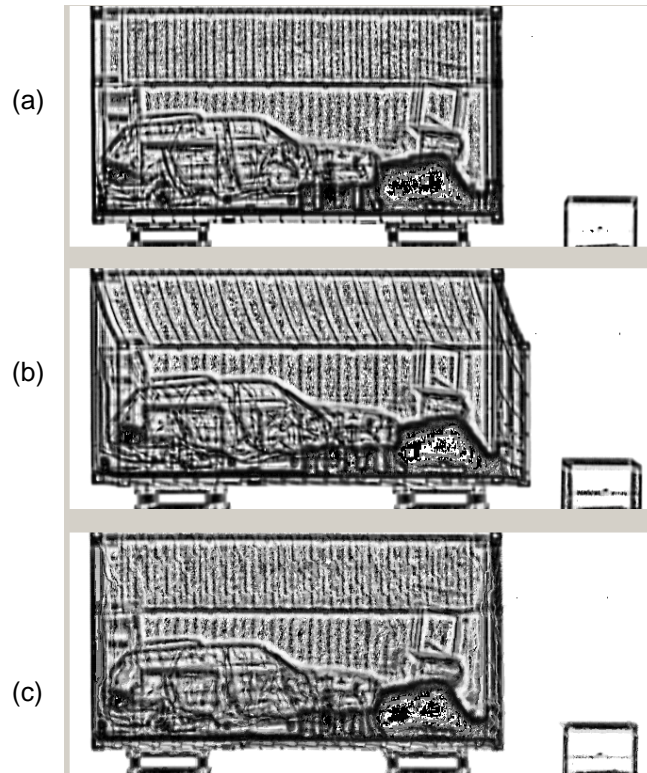

(a)

(b)

(c)

Figure 4. Contrast enhancement by adaptive window min-max. (a) Enhanced image at 0 degrees, (b) enhanced image at 10 degree and (c) deformed second image after registration.

The whole registration process on a stereo pair of 621x256 images takes less than 10 seconds on a Pentium-M 1.5GHz laptop. Once it is done, the displacement vectors for all points on the reference image are obtained. Areas with high contrasts (such as object boundaries) have more accurate matching results, whereas areas with less contrast tend to have smaller offsets (i.e., less deformation) in the estimated results than they should be. Figure 3 shows the result of warping the target image (Figure 2b) to the corresponding reference image (Figure 2a) using the estimated displacement vector field. The performance of the stereo matching algorithm,

particularly at locations with sufficient contrasts, can be seen by comparing Figure 3 and Figure 2a. However, both images have smooth intensity changes in areas such as the back of the cargo container and the roof. Therefore these areas tend not to "deform" much in the registration (i.e., deformation). Only areas with high contrasts (e.g., object surfaces that are unblocked) would provide the energy in Eq. (11) to deform pixels effectively to provide correct stereo displacements.

Studying the gamma-ray/X-Ray imaging principle will be helpful in better understanding the problem in stereo matching. Penetration of gamma-ray /X-ray can be described by the following equation:

$$I^{out} = I^{in} e^{-\mu dx} \tag{14}$$

where I is the energy strength and μ is the attenuation of the material that the ray passes through and dx is the distance between the ray penetration and exit points. The penetration of gamma-ray or X-ray energy is exponentially decaying with distance and attenuation. Thus the resulting intensity for a single point that the detector received from the two scans of different angles would vary. Therefore our stereo matching algorithm works better at those points on the boundaries of unblocked objects since the attenuation of the materials that the rays pass through would be consistent between two different views.

To partially overcome these problems, we use an adaptive window min-max method to enhance the contrast of local window and boundary. The adaptive window min-max procedure is to normalize a pixel's intensity based on the minimum and maximum intensities in a small local window of surrounding pixels. The "enhanced" image pair and the warped (deformed) image after registration are displayed in Figure 4. Comparing to the registration result without the contrast enhancement, the roof of the cargo is better registered, even though the warping artifacts due to the size of window are slightly more obvious. This could be reduced by employing the propagation approach (Yu and Bajaj 2004) in the calculation of minimum and maximum pixel values of each local window.

6.   3D ESTIMATION AND VISUALIZATION

We have also developed an interactive procedure for measuring and visualizing objects of interest. Using the interactive tool, the contour points of each object of interest are picked up by a user. For each set of points that are selected for 3D measurements and visualization, a connected 3D line-frame representation is generated and rendered in 3D.

For comparison, this interactive 3D estimation and visualization procedure is used in two ways: semi-automated stereo matching (after picking-up points) and fully-automated stereo matching for all points (before picking up points). We will discuss both of them in the following. Finally, we will compare the results obtained from multiple stereo pairs for the same target, and discuss some challenges in stereo matches of gamma-ray images.
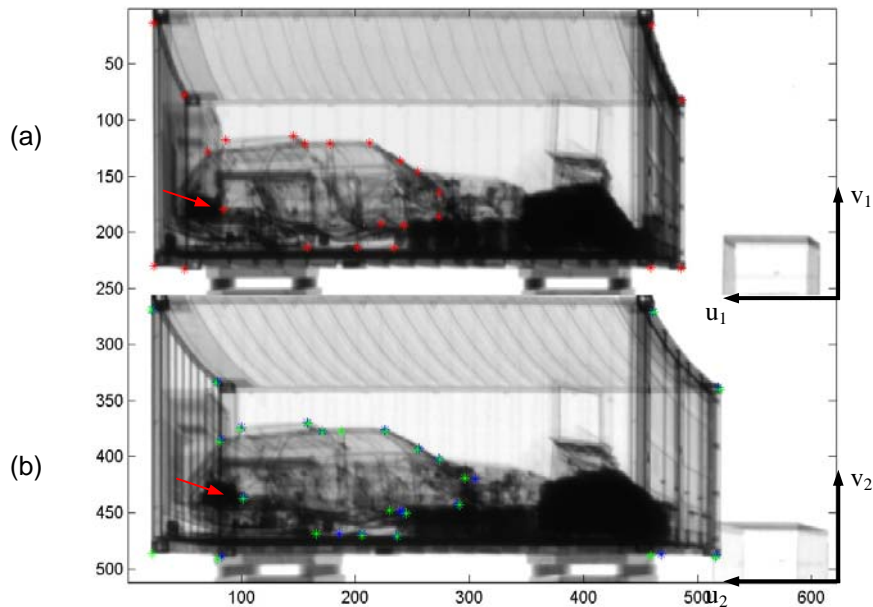

Figure 5. Pushbroom stereo matching: points selected in the top image (marked by red stars) are first searched for matches in the bottom image by a computer program (marked by blue stars), then the matches are corrected (if necessary) by a human operator for more accurate measurements (marked by green stars). Note that a few of the automated matches are incorrect.

### 6.1. Semi-automated stereo matching and 3D estimation

Our semi-automated stereo matching approach (implemented in Matlab) includes three steps: interactive point selection, correlation-based matching, and interactive matching correction. Instead of generating a dense "depth" map from a pair of gamma-ray images, we have designed an interactive user interface for selecting and measuring objects of interest. For the automated stereo matching step, we use sum of square difference (SSD) criterion on normalized images.

Figure 5 shows the process of semi-automated stereo matching for the pair of ten- and twenty-degree images. After a point in the first image is picked up by the user (marked by a red star in the first image of Figure 5), its match in the second image is automatically searched along the epipolar line of the pushbroom stereo, derived from Eq. (3). The search range is pre-determined from Eq. (4) by using the knowledge that all the objects are within the cargo container. The size of the correlation window can be determined by the user interactively. We have tried different window sizes (3x3, 9x9, 11x11, etc.) and found that 11x11 was the best for this example. The automated matches are marked by blue stars in the second image of Figure 5.

After the program finds the automated matching points, the user could correct the match if necessary (marked by green stars in the second image of Figure 5). In Figure 5, most of the automated matches are "considered" to be correct where the green marks completely overlap the blue marks. The points that are considered incorrect are those whose matches could be identified by human eyes but whose appearances are quite different between two images for automated matching. On the other hand, a few point matches that are considered to be "correct" might be incorrect; but we have no way to correct them due to the large differences between two views (e.g., the point pair pointed by arrows). In Figure 5, all eight vertices of the cargo container are selected for stereo matching as well as a few points around the boundary of a vehicle inside the cargo container. Note that the four of the eight points on the top of the container we selected here are slightly different from the ones for calibration due to the requirements of an 11x11 window centered at each point.

Together with the stereo matching interface, the reconstructed 3D structures are rendered as wire frames in 3D. For each set of points that are selected for stereo matching, a connected 3D line-frame representation is generated. Figure 6 shows several views of the 3D frame representation of the point set obtained in Figure 5: a 3D view, front view and top view. The black rectangular frame is the reconstruction of the cargo container using the calibration image data for the ten-and twenty-degree images. The red line frame is generated from the 3D measurements by the correlation-based semi-automatic stereo match algorithm. It is clearly shown that the stereo matches provide good 3D measurements for many points of the cargo container and the objects inside. Note that the selected points on the top of the container for the automated matches are
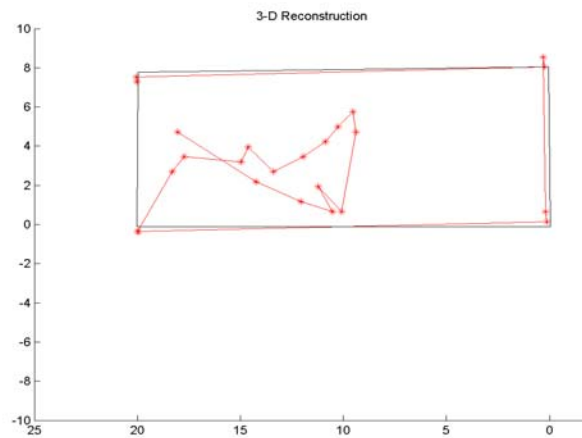
slightly off the top so that the automated matching system can find 11x11 windows in the right images. This brings in offsets between the automated 3D estimations and their calibration references.
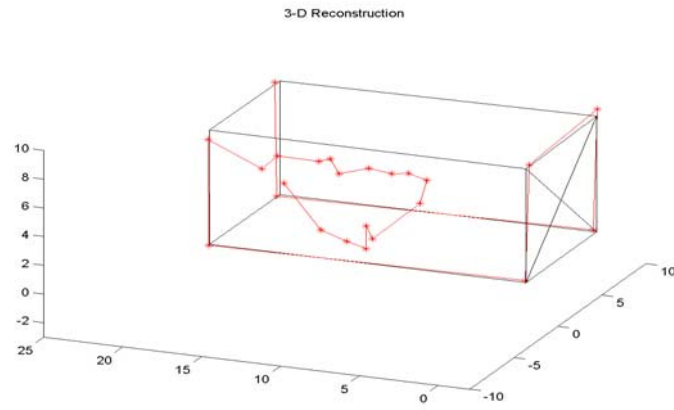
(a) 3D view
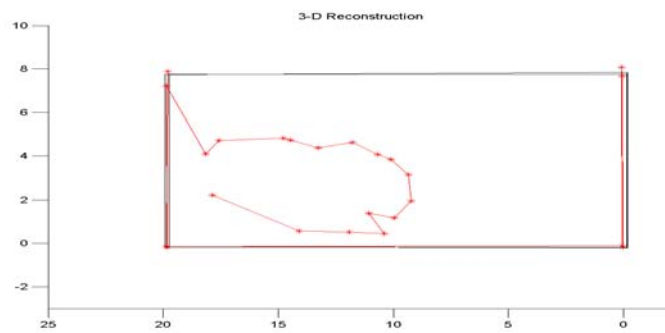
(b) Front view

(c) Top view

Figure 6. 3D measurements and visualization of objects inside the cargo container. The black rectangular frames show the cargo container constructed from the test data in Table 3. The red lines (with stars) show the 3D estimates from automated stereo matches, for the cargo container and an object inside.

We must note that the point correlation-based approach cannot guarantee correct stereo matches due to the complexity of the gamma-ray images - multiple overlapping and occlusions, repetitive patterns and low resolutions. As have shown above, a few points were matched incorrectly even with carefully selected boundary points. The need of user correction could be tedious. This is why we have developed a fully automated algorithm that utilizes some smoothness constraints. With some sacrifice to the accuracy, the automated approach is more robust, as shown in Figure 3 and Figure 4. In the following sub-section, we discuss how we use the automated stereo matching results for interactive 3D measurement and visualization.

6.2. Fully-automated stereo matching and 3D estimation

After the displacement map of a stereo pair is generated using the automated stereo matching algorithm, our interactive program (implemented in C++) allows us to pick up lists of points in the reference image, and the contour along these points in the reference image and the contour along the corresponding points in the target image are automatically drawn, side by side. In Figure 7a and Figure 7b, the pairs of colored contours show the accuracy of those points on the boundaries of objects – the roof of the cargo container, and objects inside the container. Those correspondence points are found automatically by our automated stereo matching algorithm.

Figure 7. Interactive 3D measurements. The corresponding contours are shown in (a) and (b).

The reconstructed 3D structures of those contour points that are picked up by a user in the stereo matching stage are rendered as shaded surfaces in 3D. In order to reconstruct an object, the user draws in 3D, the 2D polygon of an object contour is first triangulated in 2D. For each vertex, its 3D world coordinates are then calculated. So the 2D triangles are turned into a 3D triangular mesh. The triangulation and rendering are done with a VTK toolkit. Figure 8 shows the 3D measurements and visualization of objects inside the cargo container. The white rectangular frames show the cargo container constructed from the calibration data. The two shaded surfaces show the 3D estimates from automated stereo matches, for the roof of the cargo container and a car inside. With the 3D visualization, 3D measurements of sizes and shapes, for example, are made simple by using the most convenient views. Further object measurements and identification will be our future work.
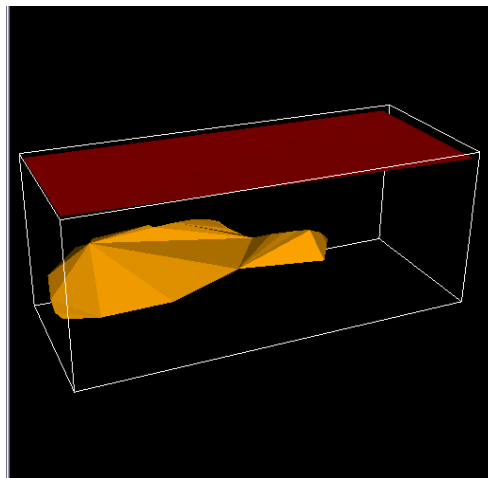

Figure 8. 3D measurements and visualization of objects inside the cargo container.

6.3. Comparison and integration of multiple stereo pairs

We have also studied how to make use of multiple pushbroom views for 3D estimation. As we have shown in Figure 2, there are three views available for the cargo container whose dimension information is known and whose imaging parameters have been obtained via the calibration method presented in Section 4. We applied the automated matching algorithm on the three views. For the purpose of comparison and integration, the ten-degree view is selected as the reference

image, and the zero- and twenty-degree views as target images. The same stereo matching procedure is applied to both pairs. Then boundaries in the reference view were picked up using our interactive program, and corresponding contours in the two target images were automatically rendered using the estimated stereo displacement vector maps.
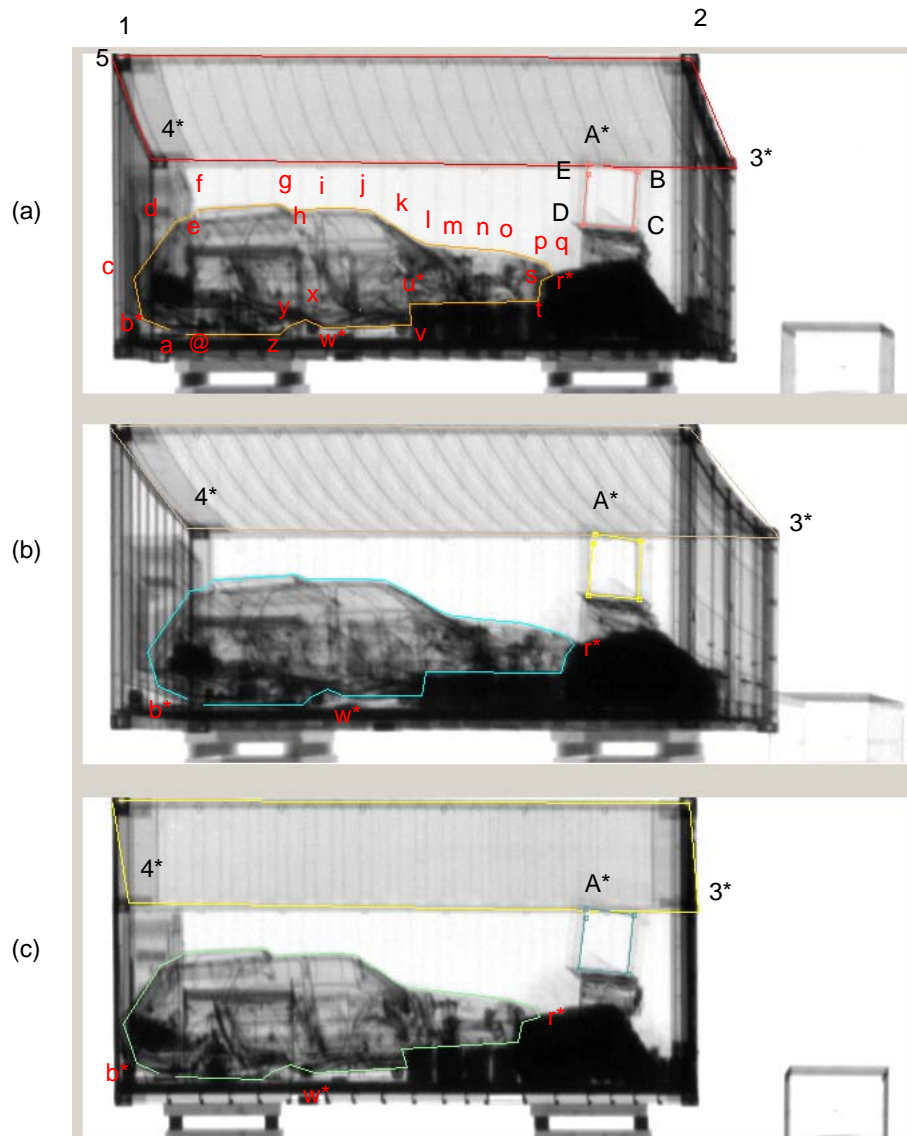

Figure 9. The corresponding contours in (a) ten-degree view (reference view), (b) twenty-degree view, and (c) zero-degree view for three objects: the cargo container roof, a car (left) and a rectangular part (right)
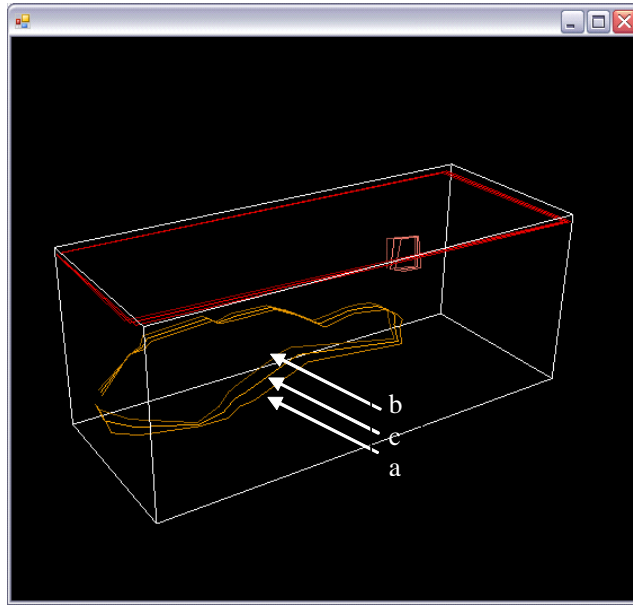
Figure 10. 3D measurements and visualization of objects inside the cargo container: (a) from the ten-zero-degree pair, (b) the ten-twenty-degree pair, and (c) their integration

Figure 9 show the three views with the user selected corresponding boundaries (of the cargo container roof, a car on the left, and a rectangular part on the right of the image), which fit quite well to the object boundaries in the two target images. Figure 10 shows the 3D estimations of the contours from the two pairs. It can be seen that the two sets of contours are consistent with each other, and the car on the left measures about 13.5 feet long and 4.0 feet high in both sets of the measurements, but there exist a small offsets. The 3D numerical results and the absolute values of the differences between the two sets of estimations are listed in Table 4. It can be seen that the differences in the x and y coordinates are very small, but the differences in the depth dimension (z) is as large as over 1 foot on average, in a container with a depth of 8 feet. Note that in theory, 1 pixel error in stereo matching will create about 0.25 feet depth error (Eq. (9)). Carefully looking into the matches reveals that the large errors in z are caused by the localization error in the u direction, due to the difficulties in finding correct matches in the gamma-ray images. This is different from the measurements in Table 3, where poor accuracy in the y direction was mainly caused by the inaccuracy in the v coordinates when picking up the images points for calibration.

Table 4. 3D measurements from multiple views: a comparison

No	From 10 & 0 degree pair			From 10 & 20 degree pair			Absolute differences (in feet)		
	x	y	z	x	y	z	|dx|	|dy|	|dz|
1	19.831	7.795	-0.049	19.846	7.844	0.041	0.015	0.049	0.090
2	0.303	7.624	-0.238	0.329	7.710	-0.078	0.026	0.086	0.160
3*	-0.007	7.406	7.267	0.164	7.768	8.302	0.171	0.362	1.034
4*	19.517	6.818	4.674	19.885	7.632	6.898	0.368	0.814	2.224
5	20.060	7.694	-0.044	20.075	7.740	0.042	0.014	0.046	0.086
a	18.505	0.811	2.145	18.744	0.914	3.589	0.239	0.104	1.444
b*	19.237	1.030	1.050	19.616	1.236	3.341	0.379	0.206	2.291
c	19.697	2.324	2.174	19.806	2.429	2.834	0.109	0.105	0.660
d	18.414	4.242	2.978	18.431	4.268	3.080	0.017	0.026	0.102
e	17.818	4.256	2.407	17.929	4.435	3.077	0.111	0.179	0.670
f	17.542	4.378	2.122	17.742	4.714	3.329	0.200	0.336	1.207
g	14.746	4.598	2.334	14.908	4.879	3.312	0.162	0.281	0.979
h	14.472	4.513	2.605	14.588	4.709	3.310	0.117	0.196	0.705
i	12.545	4.306	1.724	12.806	4.748	3.299	0.261	0.442	1.575
j	11.721	4.417	2.261	11.892	4.704	3.294	0.171	0.288	1.033
k	11.126	4.197	2.803	11.332	4.519	4.052	0.207	0.322	1.249
l	10.392	3.739	2.786	10.643	4.092	4.302	0.251	0.353	1.516
m	9.888	3.550	3.052	10.095	3.823	4.299	0.206	0.273	1.247
n	8.283	3.299	2.457	8.460	3.526	3.527	0.177	0.227	1.070
o	7.091	3.155	2.150	7.317	3.438	3.520	0.227	0.284	1.369
p	6.128	2.966	2.406	6.312	3.180	3.514	0.183	0.214	1.108
q	5.578	2.751	2.114	5.851	3.055	3.765	0.273	0.304	1.651
r*	5.441	2.479	2.389	5.753	2.792	4.272	0.312	0.312	1.883
s	6.038	2.452	3.516	6.290	2.687	5.037	0.252	0.235	1.521
t	6.176	1.843	3.798	6.297	1.931	4.529	0.121	0.088	0.731
u	10.161	1.478	1.389	10.475	1.697	3.285	0.314	0.219	1.896
v	9.976	0.804	0.550	10.261	0.939	2.268	0.284	0.135	1.718
w*	13.141	0.820	1.739	13.484	0.974	3.811	0.343	0.153	2.073
x	13.968	1.152	2.871	14.208	1.280	4.324	0.240	0.127	1.452
y	14.610	0.968	3.165	14.802	1.057	4.327	0.192	0.089	1.162
z	14.931	0.733	3.451	15.119	0.804	4.583	0.187	0.071	1.132
@	17.909	0.651	2.131	18.150	0.742	3.586	0.241	0.091	1.455
A*	3.879	5.015	0.126	4.101	5.493	1.468	0.222	0.479	1.343
B	2.229	4.922	0.365	2.284	5.036	0.695	0.055	0.114	0.331
C	2.413	3.460	0.647	2.505	3.597	1.205	0.092	0.137	0.558
D	4.109	3.508	0.409	4.242	3.712	1.215	0.133	0.204	0.806
E	3.879	4.769	0.126	4.059	5.140	1.214	0.180	0.371	1.088
Average							**0.191**	**0.225**	**1.152**

In both Figure 9 and Table 4, the "control" points (vertices) along the three polygonal boundaries are marked, and some of those with large errors are labeled with asterisks (*s).   For the contour of the cargo container roof (points #1 to #5), the three points (#1, #2 and #5) in the front match very well, but the two points (#3 and #4) at the back have matching errors. While the point pair #3 matches well in the pair of ten-twenty degree views, the point pair in the pair of ten-zero degree views is not, due to the fact that point #3 is occluded in the zero-degree view. For the rectangular part on the right (points #A - #E), the matches for the point #E is less accurate since it is mixed up with the roof pattern. The object on the right (points #a-#z) is more challenging due to its complicated structure and multiple overlapping and occlusions. For example, the point #b is corrupted by the side pattern of the cargo and a high box behind it. The point pair #r matches well in the stereo pair of ten-twenty degree views, but the point pair in the pair of ten-zero degree views is not, due to the fact that point #r is mixed with the black object on its right in the zero-degree view. Difficulties also arise for points #u and #w, for example. Nevertheless, the inaccurate depth information helps to put object into different depth layers, and also helps to get corrected *xy* measurements, which otherwise would be impossible.

To reduce the errors in depth estimation, a simple approach is to use the average of the two sets of estimations. In Figure 10, the contours integrating the two sets of measurements in such a way are also shown. A more robust approach will be to check the consistencies between two measurements for a given point from the two stereo pairs. If the difference in measurement is above a tolerable range, the measurement may not be correct.

To fully recover 3D information from gamma or x-ray images in cargo inspection is essentially a reconstruction problem from projections as in discrete tomography. However, due to limited views (3 in our example) that could be obtained in practical applications and cluttered objects inside, the problem is very much ill-posed. Nevertheless, our study also show that the proposed pushbroom stereo approach has potentials to be used for real applications when automated stereo matching processing is integrated with proper levels of user interaction. An experienced cargo inspector could pick up critical points that have obvious visual marks that enable accurate 3D measurements.

7. CONCLUSIONS AND DISCUSSIONS

In this paper we presented a practical approach for 3D measurements in gamma-ray (or X-ray) cargo inspection. The linear pushbroom sensor model is used for the gamma-ray scanning system. Thanks to the constraints of the real scanning system, we model the system by using a linear pushbroom model with only one rotation angle instead of three. This greatly simplifies the calibration procedure and increases the robustness of the parameters estimation. Using only the knowledge of the dimensions of the cargo container, we can automatically calibrate the sensor and find all the sensor parameters, including the image center, the focal length, the 3D sensor starting location, the viewing direction, and the scanning speed. The sensor modeling and calibration method is sufficiently accurate for 3D measurements.

Then, a fast and automated stereo matching algorithm based on the free-form deformable registration approach is proposed to obtain 3D measurements of objects inside the cargo. With both the automatic matching procedure and the interactive 3D visualization procedure, we hope that the 3D measurements for cargo inspection could be put into practical use.

We have made the first attempt to use pushbroom stereo for 3D gamma-ray/x-ray cargo inspection. We want to pursue this research in two directions. First, we are actively seeking collaborations with cargo inspection vendors to conduct more tests on real data in real deployments. By doing this we will obtain much important information that was not available, e.g. the real parameters of the sensor setting, the ground truth data for both sensor calibration and 3D estimation of objects under inspection, for fully evaluating our approach. Second, we will continue our study on gamma-ray stereo matching algorithms. Most of the algorithms in literature of stereo vision work well only for normal visible images. We have adopted the deformable registration method originally developed for medical imaging applications to this new application, and have incorporated adaptive contrast enhancement and epipolar geometry constraint to improve the performance of stereo matching. However, more work needs to be done to obtain dense and more accurate 3D information, particularly for small, concealed 3D objects. The knowledge of physics and optics in generating the radiographic images could be very helpful in advancing this direction of research.

ACKNOWLEDGEMENTS

REFERENCES

1. Chai, J. and H –Y. Shum, 2000. Parallel projections for stereo reconstruction. In *Proc. CVPR'00*: II 493-500.

2. Dickson, P., J. Li, Z. Zhu, A. Hanson, E. Riseman, H. Sabrin, H. Schultz and G. Whitten, 2002. Mosaic generation for under-vehicle inspection. *IEEE Workshop on Applications of Computer Vision*, Orlando, Florida, Dec 3-4

3. Gupta, R. and R. Hartley, 1997. Linear pushbroom cameras, *IEEE Trans PAMI*, 19(9), Sep.: 963-975

4. Gupta, R., A. Noble, R. Hartley, J. Mundy, A. Schmitz, 1995. Camera calibration for 2.5-D X-ray metrology. In *Proc. ICIP'95*, Vol. 3, Oct 23 - 26, Washington D.C.

5. Hardin, W., 2002. Cargo inspection: imaging solutions wait for government's call, *Machine Vision Online*, Dec.

6. Hardin, W., 2004. US Seaports: Finding the needle in hundreds of haystacks, *Machine Vision Online*, June.

7. Hitachi, 2004. Cargo container X-ray inspection systems, *Hitachi Review*, 53(2) June: 97-102. http://www.hitachi.com/rev/field/industriasystems/2006638_12876.html

8. Huang, F., S. K. Wei and R. Klette, 2006. Rotating line cameras: epipolar geometry and spatial sampling. www.ima.umn.edu/preprints/mar2006/2105.pdf

9. Keener, J. P., 1998. *Principles of Applied Mathematics: Transformation and Approximation*, Reading, MA: Addison-Wesley.

10. Klette, R., G. Gimel'farb, R. Reulke, 2001. Wide-angle image acquisition, analysis and visualization. *Proc. 14th Int. Conf. `Vision Interface (VI'2001)*, Ottawa, Canada, June, 114-125.

11. Koschan., D. Page, J.-C. Ng, M. Abidi, D. Gorsich, and G. Gerhart, 2004. SAFER under vehicle inspection through video mosaic building," *International Journal of Industrial Robot*, September, 31(5): 435-442.

12. Lu, W., M. L. Chen, G. H. Olivera, K. J. Ruchala, and T.R. Mackie, 2004. Fast free-form deformable registration via calculus of variations. *Physics in Medicine and Biology*, (49) June, pp 3067-3087

13. Noble, A., R. Hartley, J. Mundy and J. Farley. X-Ray metrology for quality assurance, In *Proc IEEE ICRA'94*, vol 2, pp 1113-1119

14. Peleg S, M, Ben-Ezra and Y. Pritch, 2001. Omnistereo: panoramic stereo imaging, *IEEE Trans. PAMI*, 23(3): 279-290.

15. Shum, H.-Y. and R. Szeliski, 1999. Stereo reconstruction from multiperspective panoramas. In *Proc. ICCV'99*: 14-21

16. Orphan, V. J., R. Richardson and D. W. Bowlin, 2002. VACIS^{TM} – a safe, reliable and cost-effective cargo inspection technology, *Port Technology International*, p. 61-65. www.porttechnology.org/journals/ed16/section02.shtml

17. Xu, C., 2000. Deformable models with application to human cerebral cortex reconstruction in magnetic resonance images. PhD Thesis, John Hopkins University.

18. Zheng, J. Y. and S. Tsuji, 1992. Panoramic representation for route recognition by a mobile robot. *International Journal of Computer Vision*, 9(1): 55-76.

19. Zhu, Z. and A. R. Hanson, 2004. LAMP: 3D Layered, Adaptive-resolution and Multi-perspective Panorama - a New Scene Representation, *Computer Vision and Image Understanding*, 96(3), Dec, pp 294-326.

20. Zhu, Z., E. M. Riseman and A. R. Hanson, 2001. Parallel-perspective stereo mosaics. In *Proc. ICCV'01*, vol I: 345-352.

21. Zhu, Z., E. M. Riseman and A. R. Hanson, 2004. Generalized parallel-perspective stereo mosaics from airborne videos, *IEEE Trans PAMI*, 26(2), Feb, pp 226-237

22. Zhu, Z, L. Zhao, J. Lei, 2005. 3D Measurements in cargo inspection with a gamma-ray linear pushbroom stereo system, *IEEE Workshop on Advanced 3D Imaging for Safety and Security*, June 25, San Diego, CA, USA, http://doi.ieeecomputersociety.org/10.1109/CVPR.2005.380

23. Zhu, Z. and Y.-C. Hu, 2007. Stereo Matching and 3D Visualization for Gamma-Ray Cargo Inspection, *Proceedings of the Eighth IEEE Workshop on Applications of Computer Vision*, Feb 21st-22nd, Austin, Texas, USA
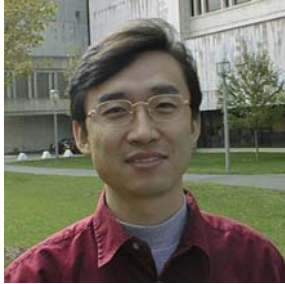
Zhigang Zhu received his B.E., M.E. and Ph.D. degrees, all in computer science from Tsinghua University, Beijing, China, in 1988, 1991 and 1997, respectively. He is currently a full professor in the Department of Computer Science, the City College and the Graduate Center at the City University of New York. He is Director of the City College Visual Computing Laboratory (CcvcL), and Co-Director of the Center for Perceptual Robotics, Intelligent Sensors and Machines (PRISM) at CCNY. Previously he has been Associate Professor at Tsinghua University, and Senior Research Fellow at the University of Massachusetts, Amherst. His research interests include 3D computer vision, Human-Computer Interaction (HCI), virtual / augmented reality, video representation, and various applications in education, environment, robotics, surveillance and transportation. He has published over 100 technical papers in the related fields. Dr. Zhu is a co-guest editor of the special issue on Modeling and Representations of Large Scale 3D Scenes, *International Journal of Computer Vision*. He is a senior member of the IEEE, a senior member of the ACM and an associate editor of the *Machine Vision Applications* Journal.

Yu-Chi Hu is a senior programmer/analyst of Department of Medical Physics at Memorial Sloan-Kettering Cancer Center and a PhD student in Computer Science at the Graduate Center at the City University of New York. He received a Computer Systems Engineer degree from Columbia University in 2003, M.S. in Computer Science from State University of New York at Binghamton in 1994, and B.B.A. in Computer and Information Science from Soochow University in Taiwan in 1989.

Li Zhao received her M.S. degree in Computer Science from The City College of New York in 2005, and B.S. in Computer Science and Technology from Beijing Institute of Meteorology in China in 1997. She is now a PhD student in Economics at the Graduate Center at the City University of New York.