Stereo Matching and 3D Visualization for Gamma-Ray Cargo Inspection

Zhigang Zhu[*ab], Yu-Chi Hu[bc]

[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*
** Email: zhu@cs.ccny.cuny.edu*

Abstract

In this paper, stereo matching and 3D visualization issues are studied for a linear pushbroom stereo model built for 3D gamma-ray (or x-ray) cargo inspection. A fast and automated stereo matching algorithm based on free-form deformable registration is developed to obtain 3D measurements of the objects under inspection. For 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. A user interface is designed for interactive 3D measurements and visualization of the objects in interests. Experimental results are presented for real gamma-ray images of a 3D cargo container and the objects inside. The 3D measurements could add more value to today's cargo inspection techniques.

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

1. Introduction

With the ongoing development of international trade, cargo inspection becomes more and more important. Quite a few X-ray or gamma (γ)-ray cargo inspection systems have been put into practical uses [5,6,7,15]. In this paper, a non-intrusive gamma-ray imaging system [15] will be used as an example to introduce our 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 practice, however, cargo inspection systems have only had two-dimensional capabilities, and human operators made most of the 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 in [3], accurate 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 online reports on cargo inspection [5,6,7].

Pushbroom images (or mosaics, when generated from video sequences) with parallel-perspective projections are very suitable for a surveillance and/or security application where the motion of the sensor used has a dominant translational direction. Examples include satellite pushbroom imaging [3], airborne video surveillance [20], 3D reconstruction for image-based rendering [1], road scene representations [18,19], under-vehicle inspection [2,10], and 3D measurements of industrial parts by an X-ray scanning system [4,12]. An advantageous feature of the pushbroom stereo is that depth resolution is independent of depth [1]. Therefore, better depth resolution could be achieved than with perspective stereo or the recently developed multi-perspective stereo with circular projection [9,13,14], given the same image resolution.

In this paper, issues on 3D measurements using a linear pushbroom stereo system are studied for gamma-ray cargo inspection. The closest work to ours is the x-ray metrology for industrial quality assurance [4,12]. To our knowledge, our research is 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) [15]. However, the algorithms developed in this paper can be used for pushbroom images acquired by X-ray or other imaging approaches as well.

In our previous work [21], 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. This paper presents a new automated stereo matching method modified from a free-form deformable registration approach [11], which is more suitable for gamma-ray images.

Fast implementation and image enhancement are also incorporated in stereo matching. In addition, we also design a user interface with interactive 3D measurement and visualization procedures that could be used in real cargo inspection applications.

The rest of the paper is organized as follows. Section 2 reviews the geometry of the pushbroom stereo model for cargo inspection. In Section 3, a fast and automated stereo matching algorithm is developed considering both the radiographic and geometric properties of gamma-ray stereo images. In Section 4, we describe a user interface designed to visualize the 3D measurements of objects in interests. Finally, in Section 5, we conclude our work and discuss a few future directions in both research and applications.

2. γ-Ray Linear Pushbroom Stereo

The system diagram of the gamma-ray cargo inspection system [15] 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.

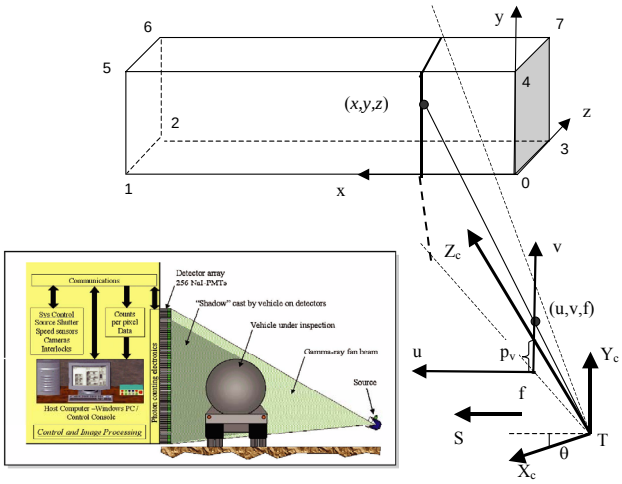

Figure 1. A gamma-ray cargo inspection system that can be modeled by linear pushbroom geometry is show in the bottom-left corner of the figure (Courtesy SAIC, San Diego, CA, USA).

The geometric model 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 camera's "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 there are no tilt and roll angles between the two coordinate systems. The sensor moves at a constant speed S (feet per scan) in the direction of the x-axis, so 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 in [3], we have the relationship between a 3D point (x,y,z) in the world coordinate system and its γ-ray image point (u,v) as

$$\begin{pmatrix} u \\ wv \\ w \end{pmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & f & p_v \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{Vx} & 0 & 0 \\ -\frac{Vy}{Vx} & 1 & 0 \\ -\frac{Vz}{Vx} & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix} \left(\begin{bmatrix} x \\ y \\ z \end{bmatrix} - \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} \right) \quad (1)$$

Note that the pushbroom scanning system has parallel projection in the u direction, but has perspective geometry in the v direction.

A dual-scanning system is a *linear pushbroom stereovision system*. It can be constructed with two approaches: two spontaneous 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 captured by such a linear pushbroom stereo system can be represented by

$$u_k = \frac{x - T_{xk} - (z - T_{zk})\tan\theta_k}{S_k} \quad (k=1, 2) \qquad (2)$$

$$v_k = f_k \cos\theta_k \frac{y - T_{yk}}{z - T_{zk}} + p_{vk}$$

where the subscript k represent the kth scan. Therefore the depth of the point can be recovered as

$$z = \frac{d - d_0}{\tan\theta_1 - \tan\theta_2} \qquad (3)$$

where

$$d = S_2u_2 - S_1u_1 \qquad (4)$$

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

$$d_0 = (T_{x1} - T_{z1}\tan\theta_1) - (T_{x2} - T_{z2}\tan\theta_2) \qquad (5)$$

is the fixed offset between two images.

Figure 2 shows two real gamma-ray images, with two different scanning angles – zero and ten degrees, respectively. Each image has a size of 621x256 pixels, i.e., 621 scans of the 256-pixel linear images. In Figure 2, the visual displacements, particularly the back surface of the cargo container, are obvious by comparing these two images. Note that Eq. (3) is acquired by only using the u coordinates of the stereo images (Eq. (4)). 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 the pushbroom stereo, the x and y coordinates of the point can be calculated from one of the two images, for example

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

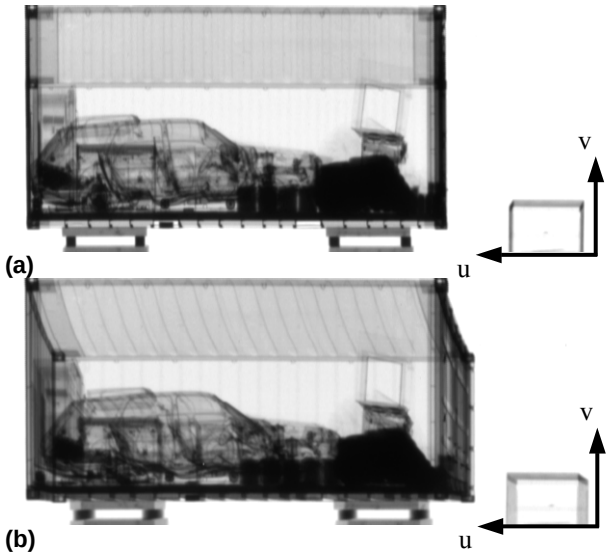

Figure 2. Real gamma-ray images with two different scanning angles (a) zero and (b) ten degrees (Courtesy SAIC, San Diego, CA, USA).

In order to use two scanning systems to calculate 3D information, we need to calibrate each scanning the 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 speech 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 $\{(x_i, y_i, z_i)\}$ and their corresponding image points $\{(u_i, v_i)\}$, i=1, 2,…, N. Our calibration method proposed in [21] only needs five known points to linearly calculate these parameters. Given the dimension of the container, which is length(x)*height(y)*depth(z)=20*8*8 (ft^3) in the example of Figure 2, we can locate at least six vertices of the rectangular container (refer to Figure 1) in each gamma-ray image by manually picking up the corresponding image points. Table 1 shows the calibration results corresponding to images (a) and (b) in Figure 2. We have two notes about the calibrations results. (1) The estimated speeds for scanning the two images are almost the same, and the angles obtained are very close to the parameters provided by SAIC, i.e., 0 and 10 degrees. (2) The two sets of focal lengths, the image centers, and the camera initial locations are consistent with each other.

Table 1. Calibrated parameters

Im	S (ft/pxl)	θ (^{o})	*F (pxl)*	p_v *(pxl)*	(T_x, T_y, T_z) *(ft)*
(a)	0.0458	0.082	427.8	21.15	(-7.419,-0.416,-14.815)
(b)	0.0457	9.399	441.2	17.79	(-9.789,-0.429,-15.141)

3. Automatic 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 [11,16] could be employed for our application. We adapt a free-form deformation registration method [11] for our gamma-ray stereo matching. There are several advantages of this method. First, it is automatic: it is a voxel based registration method (i.e., pixel based in our case of 2D images) so it does not require any feature extraction and can be done fully automatically. Second, it is fast: it is capable to find local deformation, i.e., the displacement field for each voxel (pixel), and hence a global minimization can be conducted efficiently via calculus of variations [8,16]. 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(**u**) be the reference image and B(**u**) be the target image, the displacement is defined as a mapping from A to B:

$$\mathbf{d_u} : \mathrm{A} \rightarrow \mathrm{B} \qquad (7)$$

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

$$\varepsilon(\mathbf{d_u}) = \int_{\mathbf{u}\in\Omega^2} \left[R^2(\mathbf{u}, \mathbf{d_u}) + \lambda \sum_{i=1}^{2}\sum_{j=1}^{2} \left(e_{ij}\right)^2 \right] d\mathbf{u} \qquad (8)$$

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

$e_{ij} \equiv \frac{\partial d_{u_i}}{\partial u_j}$ $(i = 1,2; j = 1,2)$ 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 **u,** $\mathbf{d_u}$ and $\frac{\partial \mathbf{d_u}}{\partial \mathbf{u}}$, by using the calculus of variations [8,16], the displacement field can be found by solving the following Euler-Lagrange equation

$$\lambda \nabla^2 \mathbf{d_u} - R(\mathbf{u}, \mathbf{d_u}) \frac{\partial R(\mathbf{u}, \mathbf{d_u})}{\partial \mathbf{d_u}} = 0 \qquad (9)$$

where $\frac{\partial R(\mathbf{u}, \mathbf{d_u})}{\partial \mathbf{d_u}}$ 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

$$d_{u\kappa}^{new} = d_{u\kappa}^{old} + \frac{L_\kappa^{old}}{\lambda + (g_\kappa^{old})^2}, \kappa = 1,2 \qquad (10)$$

Here $L_\kappa = \lambda \nabla^2 d_{u\kappa} - [B(\mathbf{u} + \mathbf{d_u}) - A(\mathbf{u})] g_\kappa (\mathbf{u} + \mathbf{d_u})$, and g_κ is the gradient in the κ's dimension $(\kappa = 1,2)$ of the image B at location $\mathbf{u} + \mathbf{d_u}$.

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 computationally efficiently at coarse level, even with low 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 number 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 our matching algorithm. 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. (2)). 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 offset of +2/-2 pixels at the vertical direction (i.e., along the y-axis) to account for geometric errors in sensor setup and sensor calibration.

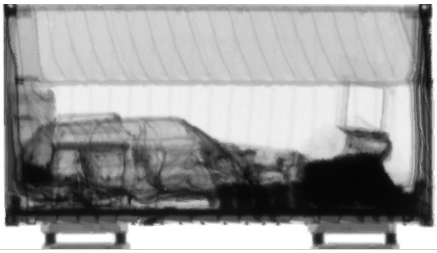

Figure 3. Warping the target images to the corresponding reference images using the estimated displacement vector fields.

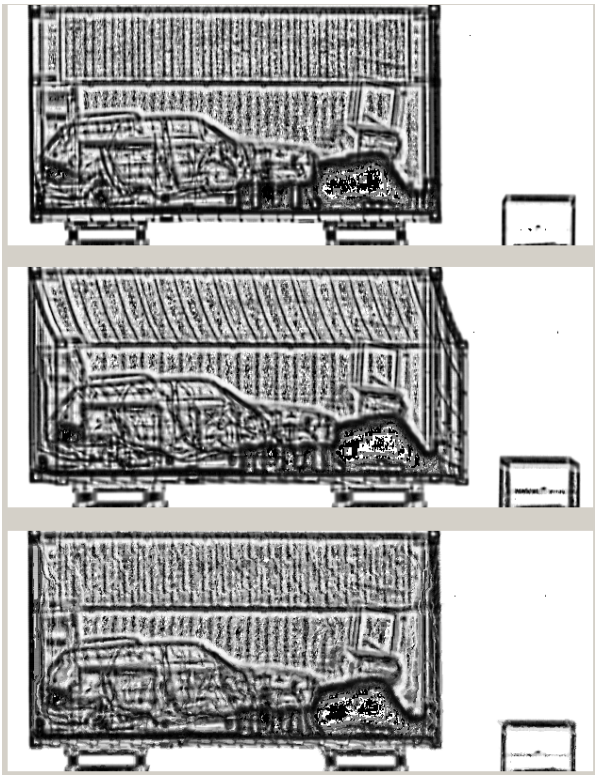

Figure 4. Contrast enhancement by adaptive window min-max. (a) Enhanced image at 0 degrees, (b) enhanced image at 10 degree and (c) deformed second image after registration.

The whole registration process 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 turn to have smaller offsets (i.e., less deformation) in the estimated results than they should be. Figure 3 shows the results of warping the target images to the corresponding reference images 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, the large area of the back of the cargo has similar intensities on both images, so does the roof area. Therefore these areas tend not to "deform" much in the registration (i.e., deformation). Only the areas with high contrasts (e.g., object surfaces that are unblocked) would provide the energy in Eq. (8) 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. The penetration of γ/X-ray energy is exponential decay 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 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 contract 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 contract 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 in [17] in the calculation of minimum and maximum pixel values of each local window.

4. 3D Measurements and Visualization

We have also developed an interactive procedure for measuring and visualizing those objects in interests. After the displacement map of a stereo pair is generated, our interactive program 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 5a and Figure 5b, 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 stereo matching algorithm.

Figure 5. Interactive 3D measurements. The corresponding pairs of contours are shown in (a) and (b).

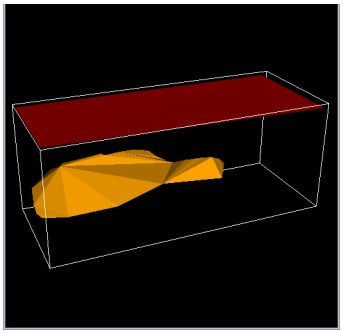

Figure 6. 3D measurements and visualization of objects inside the cargo container.

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 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 3D triangle mesh. The triangulation and rendering are done with a VTK toolkit. Figure 6 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.

5. Conclusions and Discussions

In this paper we present 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 parameter estimation. 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 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 of the 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.

6. Acknowledgments

This work is supported by AFRL under Grant Award No. FA8650-05-1-1853 and by NSF under Grant No. CNS-0551598.

7. References

[1]. Chai, J. and H –Y. Shum, 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, Mosaic generation for under-vehicle inspection. *WACV'02*.

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

[4]. Gupta, R., A. Noble, R. Hartley, J. Mundy, A. Schmitz, Camera calibration for 2.5-D X-ray metrology. *ICIP'95*.

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

[6]. Hardin, W., US seaports: finding the needle in hundreds of haystacks, *Machine Vision Online*, June 2004.

[7]. Hitachi, Cargo container X-ray inspection systems, *Hitachi Review*, 53(2) June 2004: 97-102.

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

[9]. Klette, R., G. Gimel'farb, R. Reulke, Wide-angle image acquisition, analysis and visualization. *VI'2001*, 114-125.

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

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

[12]. Noble, A., R. Hartley, J. Mundy and J. Farley. X-Ray Metrology for Quality Assurance, *ICRA'94*, vol 2, pp 1113-1119

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

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

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

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

[17]. Yu Z. and C. Bajaj, A fast and adaptive method for image contrast enhancement. *ICIP'04*, pp. 1001-1004

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

[19]. Zhu, Z., A. R. Hanson, LAMP: 3D layered, adaptive-resolution and multi-perspective panorama - a new scene representation, *CVIU*, 96(3), Dec 2004: 294-326.

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

[21]. Zhu, Z, L. Zhao, J. Lei, 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, 2005, San Diego, CA, USA