Toward a Hand-Held Laser Range Scanner: Integrating Observation-Based Motion Compensation
Hébert, P.

TOWARD A HAND-HELD LASER RANGE SCANNER: INTEGRATING OBSERVATION-BASED MOTION COMPENSATION

P. Hébert and M. Rioux

Visual Information Technology Group
National Research Council of Canada
Ottawa, Canada K1A 0R6

ABSTRACT

Although laser range sensors based on sequential scanning can provide accurate measurements in stable operation, the recovered surface geometry becomes noisy and distorted when sensors are hand-held. To compensate for camera motion, some currently existing prototypes integrate a positioning device. Unfortunately, these may not be accurate and fast enough. To circumvent this problem, a method that can compensate for motion distortion is proposed. The principle consists in using the measured shape geometry as a reference frame in 3-D space. The method is based on the collection of a redundant set of crossing profiles. Each surface profile is measured in a very short period of time such that distortion of the profile is negligible. It is assumed that the perturbation error due to motion, affects inter-profile positioning only. Then, the set of rigid crossing profiles are fitted together by moving them such as to minimize the profile intersection spacings. Experiments show that errors in the geometry can be reduced to the order of magnitude of the sensor error. The method can be integrated in the design of a hand-held sensor or as a complementary post-processing stage for improving measurement accuracy when using a sensor positioning device.

Keywords: hand-held camera, laser range sensor, motion compensation, profile registration

1. INTRODUCTION

Although laser range sensors based on sequential scanning can provide accurate measurements in stable operation, the recovered surface geometry becomes noisy and distorted when sensors are hand-held. This might appear as a limitation of sequential scanners. Nevertheless, this problem is much more general in 3-D modeling. In order to measure and build a 3-D model of an object with a single range sensor, several sets of measurements from different viewpoints must in general be gathered. The measuring process is also essentially sequential.

Using a single sensor, one usually captures a range image in stable position then moves repeatedly the object or the sensor to scan the whole surface of interest. After that process is completed, range data from different viewpoints are integrated in a common reference frame and data redundancy can be eliminated if necessary.[1–3] The integration is possible when a very accurate positioning device, such as a CMM, is used to provide the spatial relationship between the sets of measurements. It is also possible when a faster and simpler device (optical, mechanical, electromagnetic, inertial) is used, or even when the user provides manually an estimate of those spatial relationships. Though, in these last two cases, it is necessary to proceed to a global registration of the overlapping 3-D measurement sets to reduce integration error. The correction is based on the assumption that the observed shape is rigid and static during the whole scanning process.

To avoid using a positioning device and still automate the system, one must tackle the problem of correspondence between measurement sets. That problem is very difficult in the general case and has been addressed for "dense" range image data[4] as well as in passive stereo where 3-D models can be built from a sequence of images.[5] In many cases, artificial landmarks are added to facilitate correspondence[6] but this practice is not always possible or desirable.

The current challenge is the improvement of the flexibility of the scanning/integration stage in modeling. Using a hand-held sensor, one can freely move the sensor to collect 3-D measurements and reduce the acquisition and integration time. One requirement is a sensor that can collect sets of 3-D measurements in a very short period of time to reduce distortion due to hand shaking. This requirement is met by many profile range sensors such as the Biris sensor[7,8] which can collect profiles at 60 Hz. Some prototypes or commercial versions of fast active range

Other author information: (Send correspondence to P. Hébert)
P. Hébert: E-mail: hebert@iit.nrc.ca
M. Rioux: E-mail: rioux@iit.nrc.ca

image sensors have also been developed.[9–11] A second requirement is the availability of positioning devices that are compact, fast, and accurate enough compared with the sensors. This requirement seems not to be met yet. A clear example of the problem appears in results presented and discussed by Fisher *et al*[12] where an electromagnetic positioning device is used to integrate surface profiles from a hand-held sensor in a common reference frame. To reduce the error, an average filtering post-processing is proposed. Unfortunately, it significantly blurs surface details of the imaged object.

To solve the problem of automatic integration of range measurements while avoiding the necessity for an efficient positioning device, it is proposed, in this paper, to develop a profile registration method that can be applied after having collected a redundant set of crossing profiles on a surface. The method is based on an adaptation of range image registration methods to profiles.

Capturing simple patterns composed of a profile or eventually a set of profiles offers even more advantages such as allowing one to adapt scanning to the geometry of the observed object surface and avoids collecting a huge amount of redundant data during continuous scanning. The registration of a set of profiles collected by a laser range sensor is developed in the following section. Simulations as well as experimental results are presented in Sect. 3. Section 4 exposes a discussion on the results and perspectives.

2. PROFILE REGISTRATION

A profile P_i is a set of ordered range measurements in a plane. For instance, a profile may result from the intersection of a light plane with the observed surface. In the following, a profile P_i is represented as a set of coordinates, $(x_{in}, 0, z_{in})$, in the sensor reference frame where n is the index of the n^{th} profile measurement. Each surface profile is measured in a very short period of time such that distortion of the profile is negligible. Figure 1 (a) illustrates the collection of a series of profiles from different positions with respect to a global reference frame R_g. Each profile P_i is measured in the sensor reference frame R_s at position i. The spatial relationship between R_{s_i} and R_g is given by the transformation T_i. It is assumed that the perturbation error due to motion affects inter-profile positioning. Then, the set of rigid crossing profiles are fitted together by moving them such as to minimize the profile intersection spacings.

2.1. Profile Correspondence

Many algorithms and their variants have been proposed for range image registration. Actually, one should specify range image registration refinement since an initial estimate of the spatial relationship is assumed to be given. Among the algorithms that do not necessarily require feature extraction, there are the iterative closest point (ICP) algorithm,[13] its robust variants,[14] and tangent plane based variants that allow two surfaces to slide along each other.[15] Other methods in specific application conditions, and where no feature needs to be extracted have also been proposed.[16] One distinction between registration algorithms stems from the evaluation of an objective function describing the quality of the fit between the overlapping surfaces.

In the case of planar profiles in 3-D space, one must also firstly verify that the profiles overlap, or more specifically cross, on the object surface. As illustrated in Fig. 1 (b), two profiles cross if their bounded scanning planes intersect. Since the spatial relationship is not exact, two profiles P_i and P_j cross each other but do not necessarily intersect or touch at a common point. The ICP based algorithms match the closest point in an image with a given point in an overlapping reference image in order to estimate the fitting error. For two crossing profiles, the closest points on each profile are searched for. Formally, let the transformations T_i and T_j be associated with profiles P_i and P_j respectively. Each transformation can be represented by a 4×4 matrix while profile coordinates can be given in homogeneous coordinate; then the distance, d_{ij}, between two profiles, P_i and P_j, is expressed by the following expression:

$$d_{ij}(T_i, P_i, T_j, P_j) = \min_{m,n} ||T_i\ [x_{in}\ 0\ z_{in}\ 1]^t - T_j\ [x_{jm}\ 0\ z_{jm}\ 1]^t||. \tag{1}$$

To simplify the expression for d_{ij}, let X_{ij} and Y_{ij} be the crossing points of profiles P_i and P_j respectively:

$$d_{ij} = ||X_{ij} - Y_{ij}||. \tag{2}$$

Figure 1 (c) illustrates the corresponding points of two crossing profiles; in this case the points X_i and Y_j are the two closest points. Based on this correspondence, the error vector $(X_i - Y_j)$ is perpendicular to both P_i and

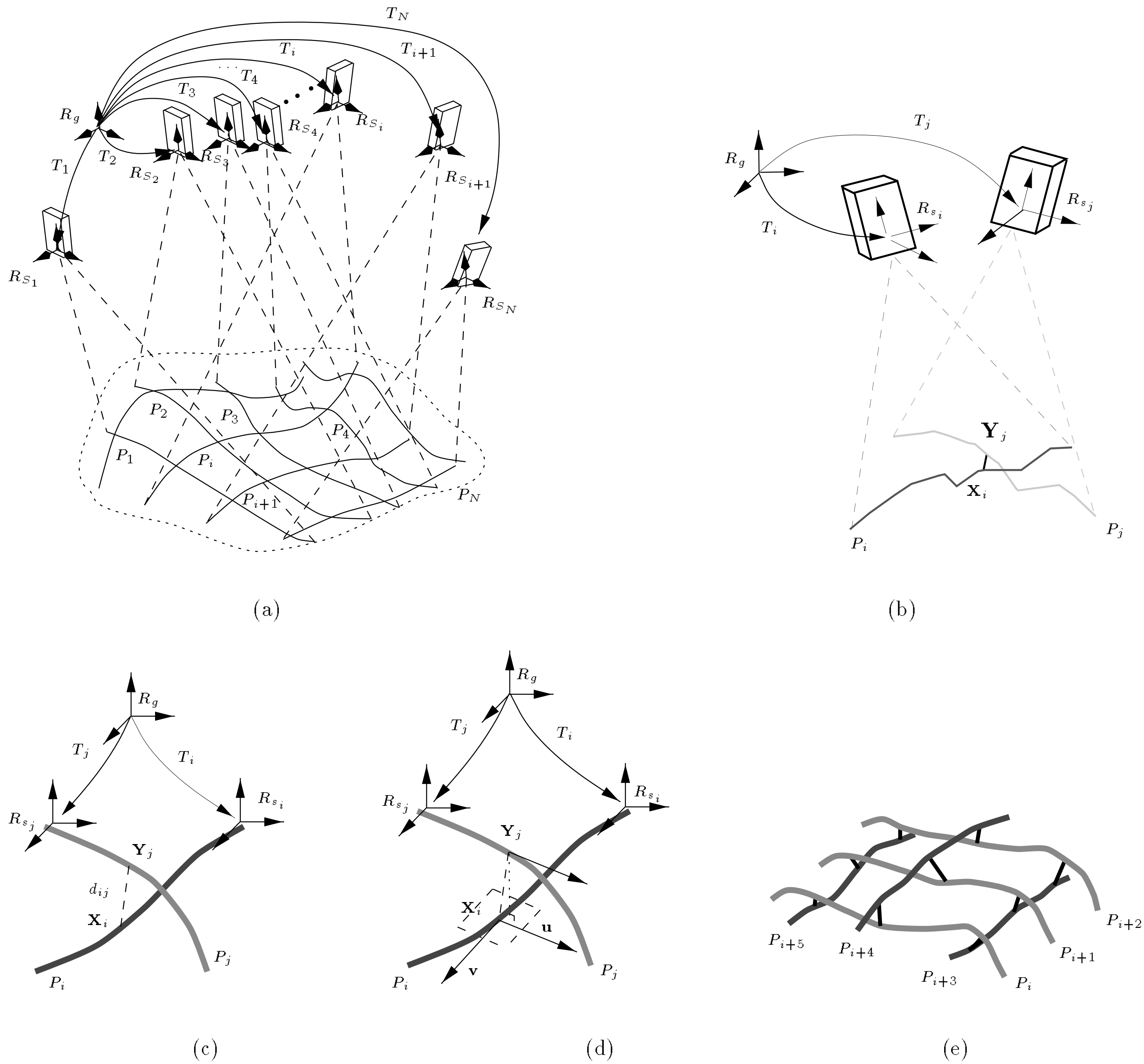


Figure 1. Profile intersections. (a) Spatial relationships between a set of profiles collected on a surface. (b) Intersection between two profiles. (c) Closest point correspondence between profiles i and j. (d) Correspondence of Y_j with the tangent plane made from tangent vectors u and v in X_i. (e) Intersecting profiles and corresponding distance vectors.

P_j when P_i and P_j are continuous (C^1) and infinite. In practice P_i and P_j are finite and can be described by a C^0 representation such as a connected set of straight line sections. For finite profiles, it is imposed that correspondence be strictly inside the boundaries to meet the crossing condition. Figure 1 (e) depicts the error vectors for a profile arrangement. Finally, the objective function can be evaluated for the whole set of crossing profiles as follows:

$$\Phi = \sum_{ij} d_{ij}^2. \tag{3}$$

The tangent plane based algorithms evaluate the quality of a fit by the distance between a point Y_j in one image j and the tangent plane evaluated at the point intersection X_i in image i resulting from intersection of the surface normal at Y_i and surface of image i. Since a surface normal cannot be evaluated on a profile, the evaluation of the objective function must be adapted to apply that type of algorithm. Figure 1 (d) illustrates the tangent plane equivalence for the intersection of two crossing profiles. For two given points X_i and Y_j on P_i and P_j respectively, the tangent vectors $\mathbf{u}$ evaluated at Y_j on P_j and $\mathbf{v}$ evaluated at X_i on P_i are used to define the tangent plane. The plane can be positioned at X_i, Y_j or in between. When X_i and Y_j are on C^1 continuous profiles, the error vector corresponding to the closest points is already aligned along the normal vector of the tangent plane. The expression for d_{ij} in Eq. 3 becomes

$$d_{ij} = A_{ij}^t Y_{ij}, \tag{4}$$

where A_{ij} is the estimated plane.

2.2. Minimizing the objective function

The objective function, Φ, estimates the lack of fit of the whole set of crossing profiles. The problem then consists in searching for the set of rigid transformations, $\{T_i\}$, that will minimize the objective function when applied to the corresponding profiles. Once again, various approaches have been proposed and applied to image registration in order to minimize this type of objective function. These include direct least squares such as in the ICP and its variants,[13,17] more robust estimations using the least median squares techniques,[14] the search over the space of admissible transformations including Levenberg-Marquardt techniques,[18] simulated annealing[19] and mechanical (dynamic) models.[20]

Usually in image registration, correspondence is reconsidered after minimization and the process is iterated. However, for the last categories of methods including Levenberg-Marquardt, simulated annealing and dynamics, correspondences are periodically updated during minimization. Many accelerators including specialized data structures have been proposed for that computationally most demanding part. The interest for using those algorithms is the possibility to simultaneously register a set of N images, or profiles in this case. However, there are convergence parameters that must be set and in the case of profiles, the large number of transformation parameters make the solution more difficult to reach in practice. Sequential update where each structure is moved while other structures are kept fixed may lead to undesirable local minima of Φ when the initial conditions are not close to the solution. Since there is a high number of profile crossings, it is proposed instead to divide an iteration into two or more stages where at each stage, a selected subset of profiles are allowed to move. The division into profile subsets can be done randomly or based on prior knowledge of the crossings.

2.3. Convergence and stability

In image registration, the iterative process can be repeated until the amount of motion is small, the value of Φ is small or, when the reduction of Φ between two consecutive iterations is below a set threshold. In practice, most of the refinement is performed in the first iterations and a maximum number of iterations is commonly imposed.

The objective function depends on the observed shape of the object and may include many local minima. It is not guaranteed that the global minimum has been reached in a given number of iterations. In the case of planar profiles, this is especially true since, when there is sensor errors, the global minimum corresponds to the trivial solution where all profiles are in a same plane. The problem is thus ill-defined unless the search is constrained in the vicinity of the initial estimates of the transformations or by adding other constraints. We propose to introduce a stabilizing cost in the objective function when the sensor's positions and orientations deviate from their original values. In this case, the objective function becomes

$$\Phi = \sum_{i,j} d_{ij}^2 + \sum_i \lambda_i ||\Delta Ti||^2. \tag{5}$$

The factor λ_i can be set empirically or based on knowledge of the sensor position error when a positioning device is used.

3. SIMULATIONS AND EXPERIMENTS

A series of simulations and experiments were performed to assess the feasibility of profile registration in different situations with different algorithm variants. The first simulation results in Fig. 2 show the compensation of translation perturbations. A 256×256 range image of various objects was captured in stable position and was used as a reference model. Figure 2 (a) illustrates the range image where closer objects appear lighter. Each pixel covers an area of approximately $0.5\,mm \times 0.5\,mm$. From this reference model, both an horizontal and a vertical image of perturbed profiles is generated. Each original profile is then replaced by a neighboring profile whose index is randomly generated and after having introduced a random offset along the two axes within the profile. Perturbed images are shown in Figs. 2 (c) and (d) where the perturbation on motion corresponds to $\sigma_x = \sigma_y = \sigma_z = 1\,mm$ (uncorrelated). Figure 2 (b) shows the ideal corrected image. Isolated black pixels show areas where the surface has not been sampled. Figure 2 (e) shows the actual corrected image after application of the tangent plane method. Since this particular case leads to a linear problem, a Cholesky decomposition was used to solve at each iteration in two steps, using horizontal and vertical profiles alternately. Figure 2 (f) shows the comparison of the result with Fig. 2 (b). Lighter areas represent regions where the error is more important. Errors within a pixel appear on edges. A part of the error arises from the absence of subpixel interpolation in the final image reconstruction. Figures 3 (a) to (f) show the same simulation for a flat object laid down on a planar surface. Although the distribution of surface orientations is more limited in this case, the compensation is still significant.

Different simulations with variants of algorithms including ICP, Levenberg-Marquardt and different minimizing schemes such as sequential correction of profiles or sets of profiles have revealed that non-linearities introduced by rotation errors may require a closer initial solution to converge properly. To implement the proposed stabilizer in Eq. 5, the initial sensor positions and orientations were associated with points along their reference frame axes and those points were matched all along the iterative process.

Learning from those simulations, experiments were conducted where the first presented experiment aimed at testing the method in severe perturbations with coarse initial estimates. The "pumpkin object" partially shown in Fig. 4 (e) follows grossly the shape of a one meter cylinder with a $20\,cm$ radius. The object was hand-held in front of a range sensor mounted on a tripod at a distance of nearly $2\,m$ from the object. Two images were collected; one image while the object was held nearly horizontal and a second image while it was held nearly vertical. During both $33\,s$ image scanning procedures, the object was shaken with an amplitude of approximately $20\,cm$. Figures 4 (c) and (g) illustrate views of the two sets of profiles. To assess the level of perturbation by comparison, Figs. 4 (d) and (h) show similar views of the object for profiles collected in stable (static) positions. The original perturbed vertical set of profiles was grossly manually aligned with the horizontal set of profiles to initiate the correction process. Figures 4 (b) and (f) show the two corrected sets of profiles after 30 iterations of the adapted ICP version. Despite strong perturbations, one can observe that the shape was significantly corrected. Figure 4 (a) shows the combination of the two corrected sets of profiles.

The next experiment shows the application of the algorithm to improve results when using a positioning device. A Biris sensor mounted on a Faro mechanical arm was used to sweep a section of a mannequin head face along two directions following the surface. Due to positioning device, sensor mounting calibration, and synchronization errors, the crossing profiles did not match perfectly. Figure 5 (a) shows the whole set of profiles in gray shade and the correspondence error vectors in dark shade before profile registration. Figure 5 (b) shows the corrected profiles and the corresponding error vectors after 20 iterations. The three curves in Fig. 6 show the monitored error, rotation and translation with respect to iterations respectively. The global squared error norm is merely the sum of the squared norm of error correspondence vectors. Actually, it corresponds to the value of the objective function in Eq. 3. After 20 iterations, the average error norm is below the positioning device error and about three times the sensor error. The global translation squared norm curve results from the monitoring of the sum of translation vector squared norms applied to every profile. Finally, the rotation is monitored by evaluating the expression,

$$\sum_{\{P_i\}} \sum_{k=1}^{3} (r_{i_{kk}} - 1)^2, \tag{6}$$

at each iteration, where $r_{i_{kk}}$ is a diagonal term of the rotation matrix associated with profile P_i. Those curves show a typical behavior of registration refinement algorithms where most corrections are performed in the very first iterations.

4. DISCUSSION

This paper has proposed a correction method based on an adaptation of surface range image registration algorithms to profiles. To build a rigid set from several crossing profiles, the main adaptations concern:

- Estimating the correspondences.
- Dividing the set of profiles to proceed in two steps at each iterations.
- Introducing a constraint to stabilize the minimization of the objective function where the sensors' positions and orientations disparities are penalized.

Several simulations and experiments were conducted to develop and evaluate the potential of profile registration. Some of the results are reported here and show the feasibility of profile registration refinement to compensate for scanning errors due to motion. Profile registration was applied in both cases when positioning devices are used or not. The results, although qualitative or globally quantitative, appear promising. It would be desirable to lower the average error closer to the sensor error level. This could be done by progressively relaxing the stabilizing constraints, but it is not obvious yet how it can be done automatically and rigorously. Local quantitative analysis and comparison with ground truth surface models will be made possible using an industrial part inspection software currently under development at our laboratory. Future developments in profile registration will integrate intensity or color in addition to shape.[21]

An important motivation under that work is the possibility that a hand-held scanner could provide to sweep the object surface in a continuous motion. To improve stability and efficiency of algorithms, one can consider using patterns composed of profiles like, for instance, a cross-hair or any more complex patterns of rigid profiles in the design of a hand-held sensor. The principle of the method remains identical.

5. ACKNOWLEDGMENTS

The authors would like to thank Étienne Moreau for his participation in the development of software, Luc Cournoyer for his technical assistance provided in the course of some experiments. The authors wish to express their gratitude to Jacques Domey, Pierre Boulanger and especially to Guy Godin for many valuable discussions and enlightened comments.

REFERENCES

1. M. Soucy and D. Laurendeau, "A general surface approach to the integration of a set of range views," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **17**(4), pp. 344–358, 1995.
2. "http://www.innovmetric.com."
3. G. Turk and M. Levoy, "Zippered polygon meshes from range images," in *proceedings of SIGGRAPH 94*, pp. 311–318, (Orlando, FL.), July 1994.
4. R. Bergevin, D. Laurendeau, and D. Poussart, "Registering range views of multipart objects," *Computer vision and image understanding* **61**(1), pp. 1–16, 1995.
5. O. Faugeras, *Three-dimensional computer vision: a geometric viewpoint*, MIT press, 1993. 663 p.
6. W. Niem and J. Wingbermühle, "Automatic reconstruction of 3D objects using a mobile monoscopic camera," in *Proceedings of the international conference on recent advances in 3-D digital imaging and modeling*, pp. 173–180, May 1997. Ottawa, Canada.

7. F. Blais, M. Rioux, and J. Domey, “A compact three-dimensional camera for robot and vehicle guidance,” *Optics and Lasers in Engineering* **10**, pp. 227–239, 1989.
8. “http://www.vitana.com.”
9. J. A. Beraldin, F. Blais, M. Rioux, J. Domey, and L. Cournoyer, “A video rate laser range camera for electronic boards inspection,” in *proceedings of the vision'90 conference*, pp. 4.1–4.11, Nov. 1990. Detroit, MI.
10. “http://www.eois.com.”
11. M. Proesmans and L. V. Gool, “One-shot active 3D image capture,” in *Three-dimensional image capture*, R. N. Ellson and J. H. Nurre, eds., vol. 3023, pp. 50–61, SPIE-The international society for optical engineering, Feb. 1997. San Jose, CA.
12. R. B. Fisher, A. Fitzgibbon, A. Gionis, M. Wright, and D. Eggert, “A hand-held optical surface scanner for environmental modeling and virtual reality,” Tech. Rep. 778, Department of artificial intelligence, University of Edinburgh, 1996.
13. P. J. Besl and N. D. McKay, “A method for registration of 3-D shapes,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **14**(2), pp. 239–256, 1992.
14. T. Masuda and N. Yokoya, “A robust method for registration and segmentation of multiple range images,” *Computer vision and image understanding* **61**(3), pp. 295–307, 1995.
15. Y. Chen and G. Medioni, “Object modelling by registration of multiple range images,” *Image and vision computing* **10**(3), pp. 145–155, 1992.
16. B. Kamgar-Parsi, J. L. Jones, and A. Rosenfeld, “Registration of multiple overlapping range images: Scenes without distinctive features,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **13**, pp. 857–871, Sept. 1991.
17. H. Gagnon, M. Soucy, R. Bergevin, and D. Laurendeau, “Registration of multiple range views for automatic 3-D model building,” in *IEEE Computer society conference on computer vision and pattern recognition*, pp. 581–586, June 1994. Seattle, WA.
18. G. Champleboux, S. Lavallee, R. Szeliski, and L. Brunie, “From accurate range imaging sensor calibration to accurate model-based 3-D object localization,” in *IEEE Computer society conference on computer vision and pattern recognition*, pp. 83–89, June 1992. Champaign, IL.
19. G. Blais and M. D. Levine, “Registering multiview range data to create 3D computer objects,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **17**(8), pp. 820–824, 1995.
20. D. W. Eggert, A. W. Fitzgibbon, and R. B. Fisher, “Simultaneous registration of multiple range views satisfying global consistency constraints for use in reverse engineering,” Tech. Rep. 804, Department of artificial intelligence, University of Edinburgh, 1996.
21. G. Godin, M. Rioux, and R. Baribeau, “Three-dimensional registration using range and intensity information,” in *Proceedings of SPIE Videometric III, Boston, 2-4 Nov.*, vol. 2350, pp. 279–290, 1994.

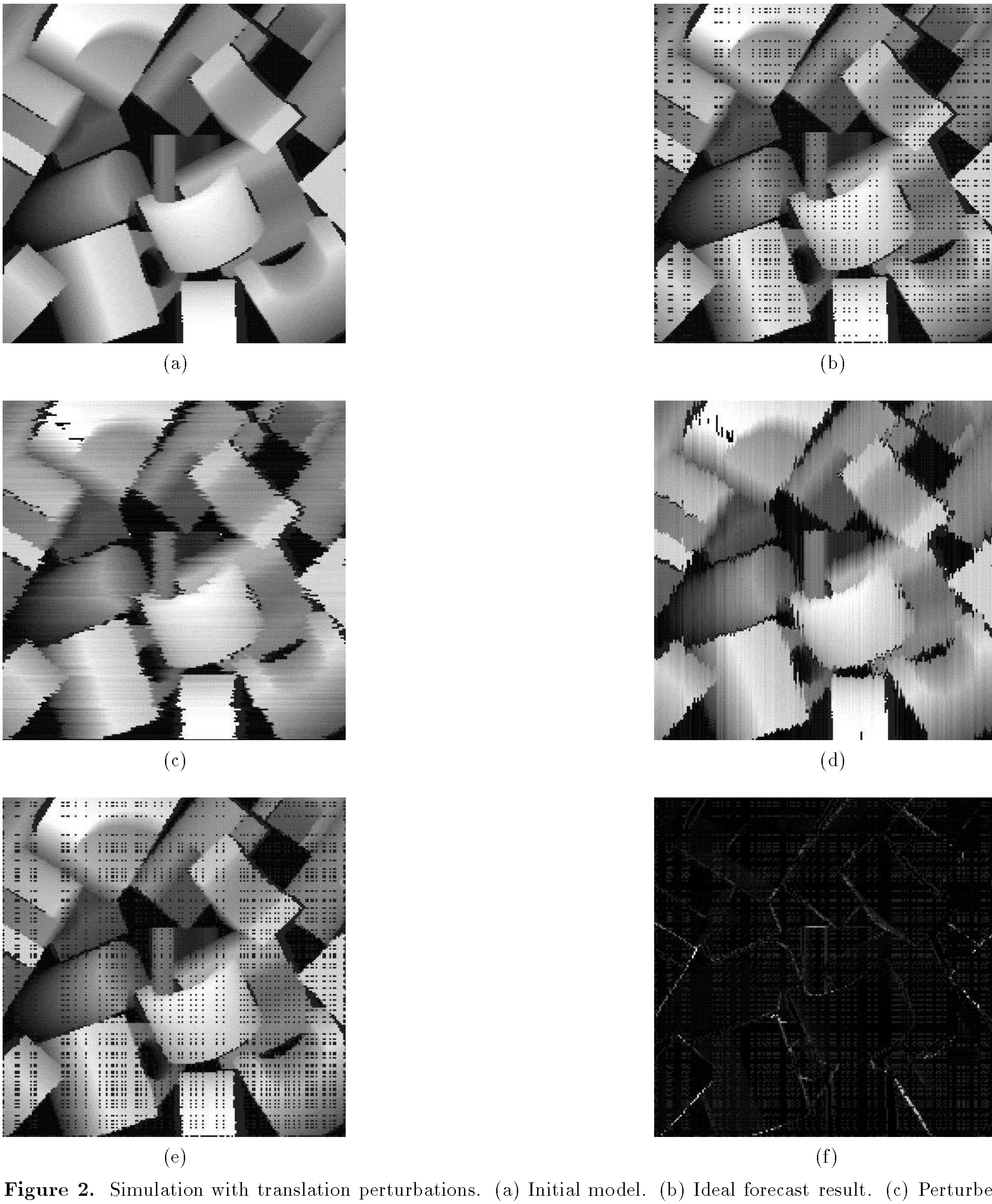


Figure 2. Simulation with translation perturbations. (a) Initial model. (b) Ideal forecast result. (c) Perturbed horizontal image. (d) Perturbed vertical image. (e) Result after compensation from horizontal and vertical images. (f) Difference with the ideal forecast image.

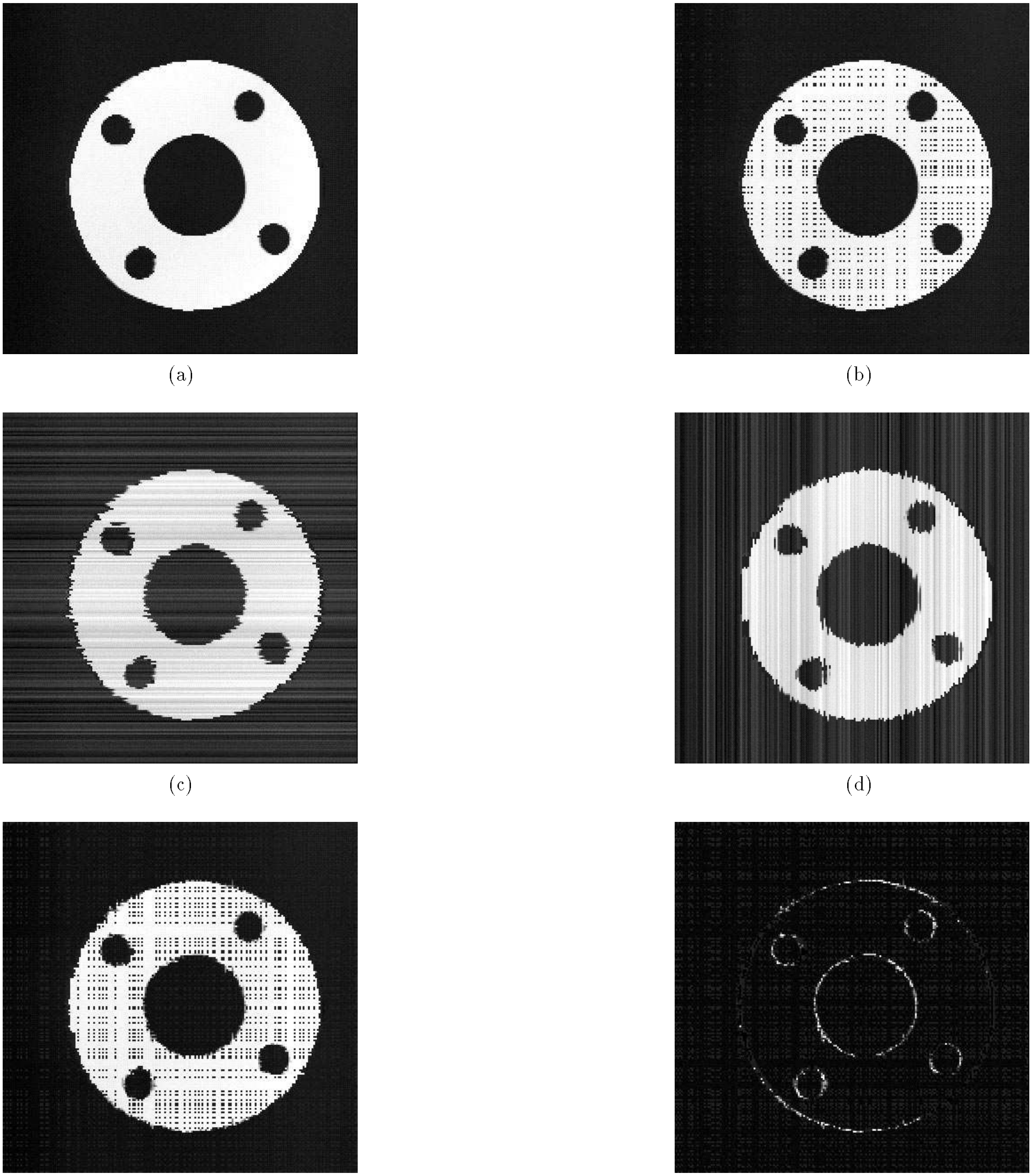


Figure 3. Simulation with translation perturbations. (a) Initial model. (b) Ideal forecast result. (c) Perturbed horizontal image. (d) Perturbed vertical image. (e) Result after compensation from horizontal and vertical images. (f) Difference with the ideal forecast image.

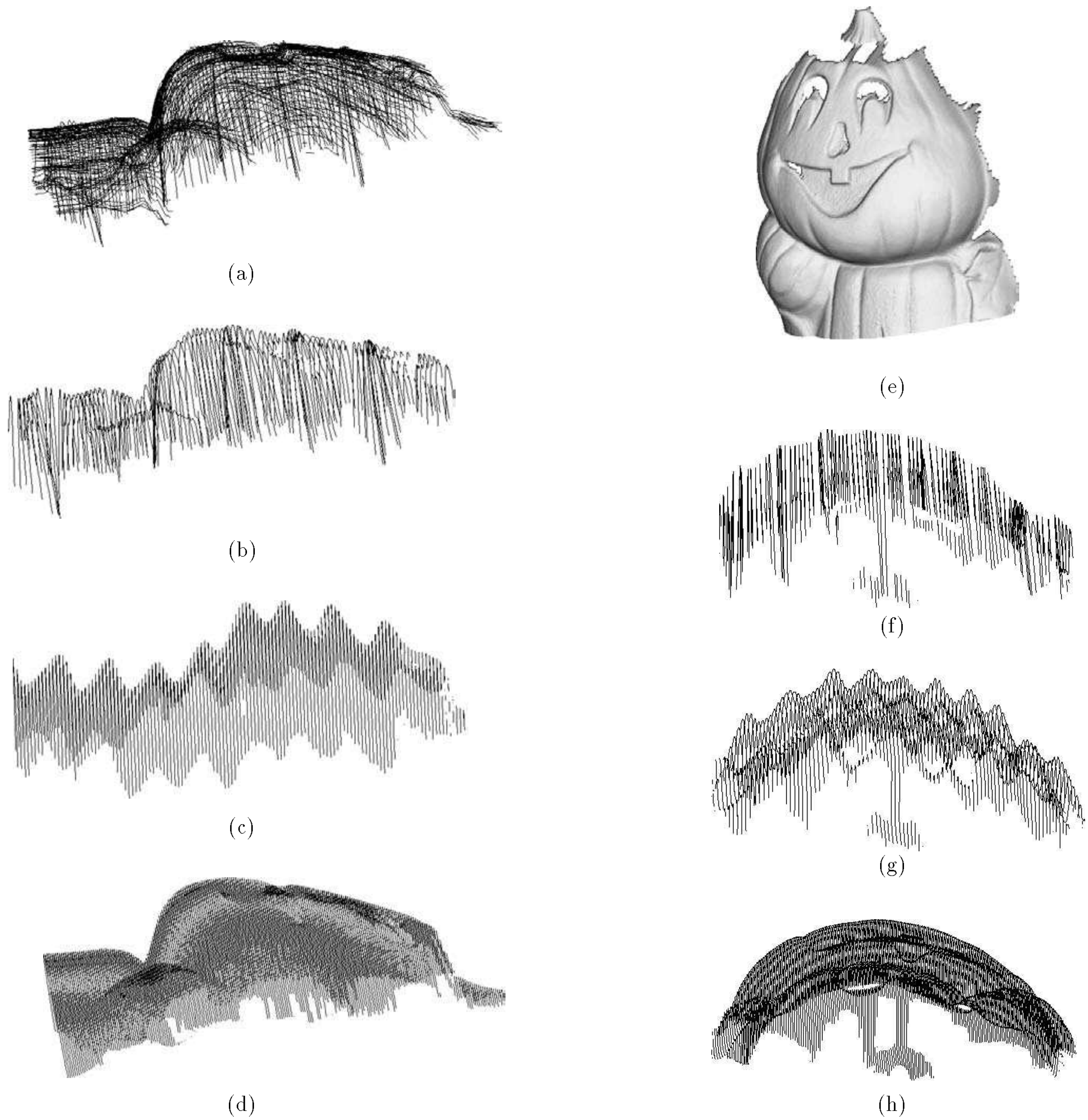
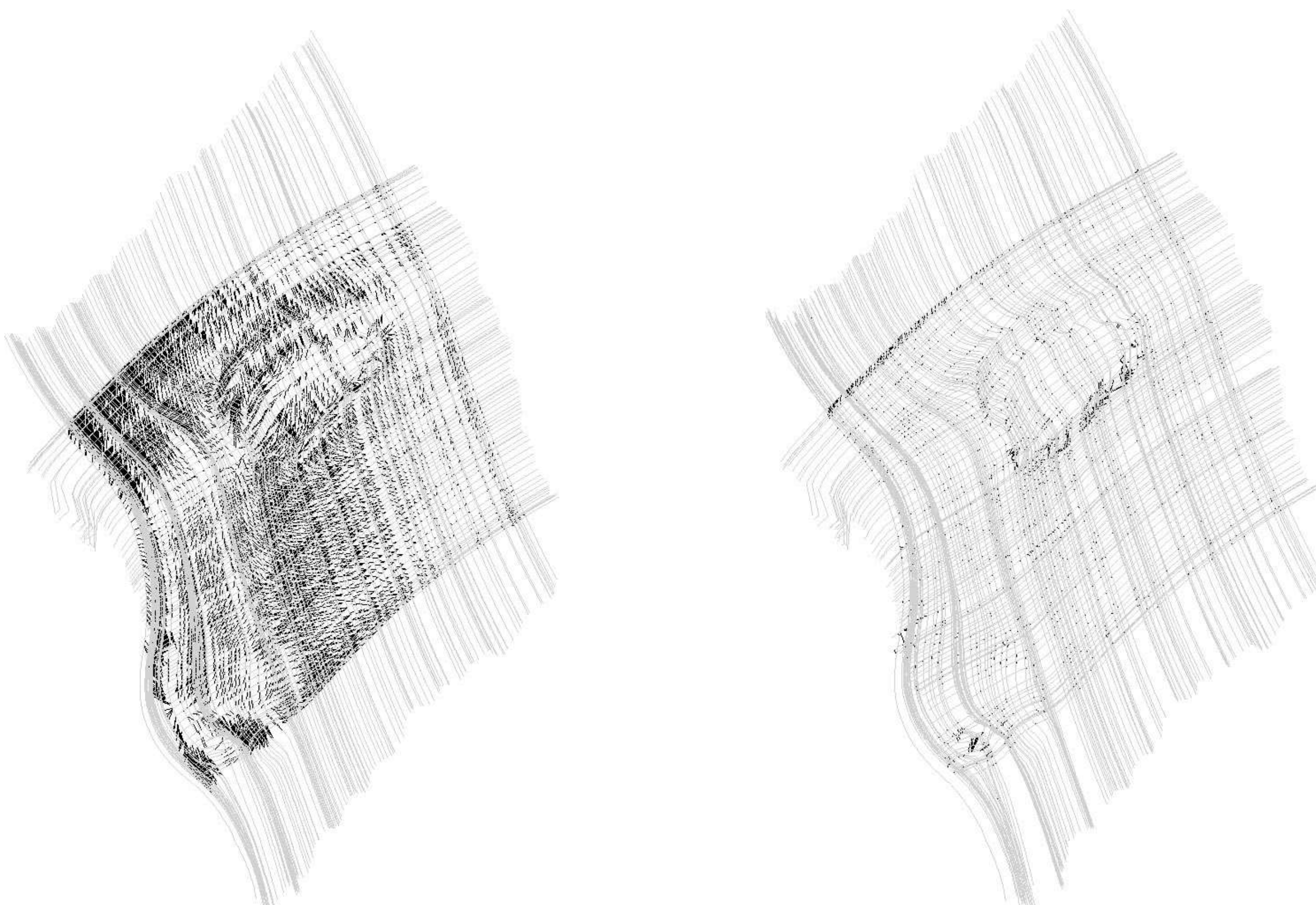
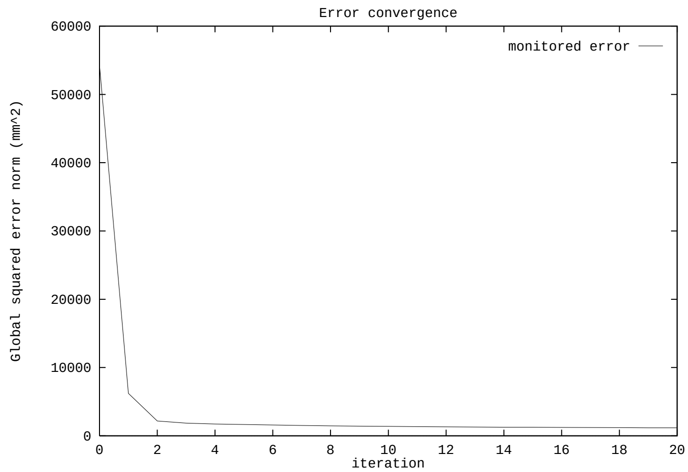


Figure 4. Experimental results with strong perturbations both in rotation and translation. (a) Compound result after correction of both horizontal and vertical acquisitions. (b) Corrected vertical acquisition. (c) Similar view (with b) of non-corrected vertical acquisition. (d) Similar view (with b) of a dense static acquisition. (e) Range image of the object captured in static position. (f) Corrected horizontal acquisition. (g) Similar view (with f) of non-corrected horizontal acquisition.(h) Similar view (with f) of a dense static acquisition.

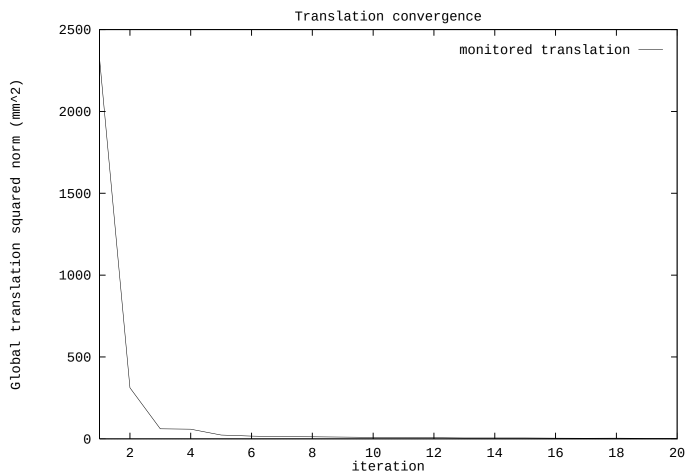


(a) (b)

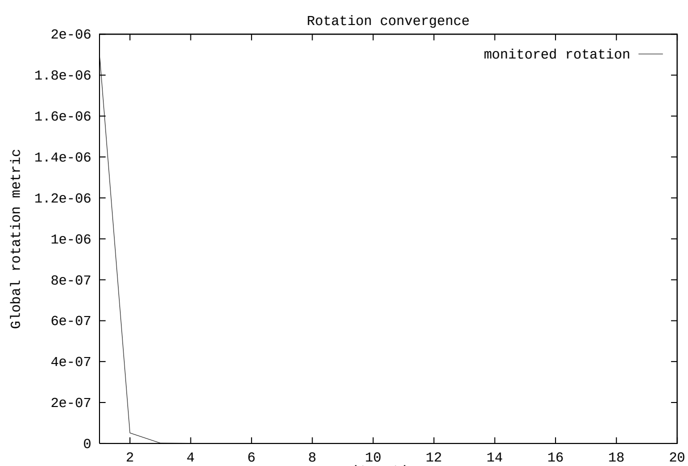
Figure 5. Positioning device error correction. (a) Two sets of profiles in gray shade with the intersecting error correspondence vectors in dark shade. (b) The same two sets of profiles after 20 iterations of correction.



(a)



(b)



(c)

Figure 6. Convergence of the correction process for the mannequin experiment. (a) Progression of the global squared error norm. (b) Progression of the global translation squared norm. (c) Progression of the global rotation metric.