



HAL
open science

An innovative hand-held vision-based digitizing system for 3D modelling

Benjamin Coudrin, Michel Devy, Jean-José Orteu, Ludovic Brèthes

► **To cite this version:**

Benjamin Coudrin, Michel Devy, Jean-José Orteu, Ludovic Brèthes. An innovative hand-held vision-based digitizing system for 3D modelling. *Optics and Lasers in Engineering*, 2011, 49 (9-10), pp.1168-1176. 10.1016/j.optlaseng.2011.05.004 . hal-01171130

HAL Id: hal-01171130

<https://imt-mines-albi.hal.science/hal-01171130>

Submitted on 3 Jul 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An innovative hand-held vision-based digitizing system for 3D modelling

Benjamin Coudrin^{a,b,c,d,*}, Michel Devy^{b,c}, Jean-José Orteu^d, Ludovic
Brèthes^a

^a*NOOMEO ; rue Galilée, BP 57267, 31672 Labège CEDEX, France*

^b*CNRS ; LAAS ; 7 avenue du colonel Roche, F-31077 Toulouse Cedex 4, France*

^c*Université de Toulouse ; UPS, INSA, INP, ISAE ; UT1, UTM, LAAS ; F-31077
Toulouse Cedex 4, France*

^d*Université de Toulouse ; Mines Albi ; Institut Clément Ader (ICA) ; Campus Jarlard,
F-81013 Albi, France*

Abstract

We describe a new hand-held 3D modelling device using vision and inertial measurements. Our system allows fast and accurate acquisition of the geometry and appearance information of any 3D object. We focused our work towards an easy manipulation and operating condition.

Our methods allow automatic registration with no preparation of the scene (i.e. no markers) even when the object is moved between two acquisitions.

In this paper, the design of the system and the developed methods for its use are presented. The system has been evaluated, qualitatively and quantitatively, using reference measurements provided by commercial scanning devices. The results show that this new hand-held scanning device is really

*Tel : (+33)5.61.00.77.18 ; Fax : (+33)9.74.76.02.52

Email addresses: benjamin.coudrin@noomeo.eu (Benjamin Coudrin),
michel.devy@laas.fr (Michel Devy), jean-jose.orteu@mines-albi.fr (Jean-José Orteu), ludovic.brethes@noomeo.eu (Ludovic Brèthes)

competitive for modelling any 3D object.

Keywords: 3D digitizing; 3D modelling; hand-held; registration

1 **1. Introduction**

2 3D Object Modelling has generated many research works in the last
3 decades, especially for Non Destructive Testing or Quality Control appli-
4 cations. Several systems can be found off-the-shelf, either based on vision [1]
5 or on laser ranging [2]. New applications where 3D modelling must be per-
6 formed by non expert people are arising, for instance for archeologists, cu-
7 rators or artists who want to model art objects. Moreover, several potential
8 mass market applications are emerging, e.g. 3D modelling from smart devices
9 either embedded on assistance robots or moved by hand around an object.

10 These new applications involve several constraints. In particular, the 3D
11 modelling task is often executed in a human environment, typically at home,
12 where it is better to avoid laser-based technologies. Moreover the system
13 must be user-friendly, lightweight to be hand-held (i.e. compact), operated
14 with a minimal preparation of the environment (i.e. no marker), and cheaper
15 than existing systems. In that context, Vision, generally associated with
16 illuminators [3, 4], is the better way to acquire sensory data. Some hand-held
17 systems exist in the industrial community. Mostly these systems require to
18 place a reference in the scene. Some use magnetic references or photometric
19 targets. This leads to a more complex digitization process and can forbid to
20 move the object during the process.

21 This paper presents a new Vision-based hand-held 3D scanner allowing
22 acquisition of geometry and texture from an object without any scene prepa-

23 ration. The object is first represented by clouds of 3D points and then with a
24 textured mesh. Digitizing 3D objects requires a sequential process of geomet-
25 ric modelling from visual data acquired from camera(s) moved by a robot or
26 by an operator around the object. 3D reconstruction gives a 3D point cloud
27 from each acquisition. Clouds of 3D points are fragmented representations of
28 space that need to be merged. This imposes to know the shooting positions
29 of the scanner. Our scanner being hand-held, we have no *a priori* model for
30 this motion, like we would have if the scanner was held by a robot or if we
31 had placed targets in the scene.

32 Several 3D registration algorithms have been proposed since a long time
33 in order to merge 3D data obtained from several viewpoints. Besl *et al.*
34 introduced the main method used to estimate the rigid transformation be-
35 tween two clouds of 3D points, the so-called ICP algorithm (*Iterative Closest*
36 *Point*) [5]. This method has since been adapted with numerous variations [6]
37 in order to speed up the convergence and decrease the complexity. Sandhu
38 *et al.* [7] recently combine ICP with a particular filtering approach. Li
39 *et al.* [8] improved the robustness, providing a guaranteed registration. Our
40 registration method combines several operations, based on IMU (Inertial Mo-
41 tion Unit) measurements, ICP-based algorithms and visual odometry from
42 tracked interest image-points.

43 The aim of this new 3D modelling system is to provide an easy to use
44 device. This means that we focused our work on providing reliable and mostly
45 automatic operation. Being hand-held, and lightweight, the device is easily
46 moved and can access a large range of view points, allowing to scan complex
47 shape objects in details or in difficult access conditions. Since no preparation

48 is required, the overall modelling process has been made shorter and the skill
49 level required for the operator has been reduced. Moreover, the system only
50 requires to be plugged to a computer, which allows to use it with a laptop
51 in outdoor scanning tasks for example.

52 Next section is devoted to the sensor description. Then section 3 presents
53 the main modelling steps: image acquisition, 3D reconstruction and 3D reg-
54 istration. In section 4, several models built from our approach are evaluated,
55 from comparisons either with the actual models for known objects, or with
56 models built with commercial scanning systems. Finally, section 5 presents
57 some outlooks for our future works.

58 **2. Description of the system**

59 This new vision-based digitizing system is principally composed of two
60 cameras used as a stereovision pair. 3D data is obtained by triangulating
61 image points between a pair of images. Matching pixels in areas of homo-
62 geneous intensity in images is impossible. To tackle this problem, we use a
63 projection of a speckle pattern to give the observed scene a suitable random
64 texture.

65 At a given instant, the system can only acquire a partial view of the scene
66 due to the field of view of the cameras. Modelling a scene from a hand-held
67 vision-based system requires to aggregate these partial views acquired at
68 several instants. This is the problem of estimating the pose or the motion of
69 the digitizing system throughout time. To help us in this task, the digitizing
70 system has been equipped with an inertial motion unit (IMU) able to measure
71 the dynamic or the attitude of the digitizing system.

72 The system weights 1.8 kg, its dimensions are $220 \times 240 \times 90$ mm. Figure
 73 1 describes our system and the frames associated with the sensors.

74 [Figure 1 about here.]

75 2.1. Cameras

76 The digitizing system is mainly built around two cameras. The cameras
 77 are equipped with a global shutter CCD sensors with a 1024×768 resolution.
 78 They can operate up to a 39 Hz maximum frame rate. However, using a frame
 79 rate close to the maximum decreases image quality.

80 The first camera, placed at the bottom of the system, is called the *coaxial*
 81 *camera* because it shares its frame with the absolute frame chosen for the
 82 system. The *coaxial frame* is noted \mathcal{C}_0 .

83 The second camera, placed at the top of the system, is the *lateral camera*.
 84 Its frame is noted \mathcal{C}_1 . We suppose that the transformation between the *lateral*
 85 and the *coaxial* frames is known¹ and noted $\mathbf{H}^{\mathcal{C}_1\mathcal{C}_0}$, such that :

$$\begin{bmatrix} p^{\mathcal{C}_1} \\ 1 \end{bmatrix} = \mathbf{H}^{\mathcal{C}_1\mathcal{C}_0} \begin{bmatrix} p^{\mathcal{C}_0} \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} p^{\mathcal{C}_0} \\ 1 \end{bmatrix} \quad (1)$$

86 where \mathbf{R} and \mathbf{t} are respectively the rotation matrix and translation vector
 87 relating the *coaxial frame* and the *lateral frame*.

88 It should be noted that, to guarantee the best precision in matching and
 89 to avoid noise, the cameras need to be perfectly synchronized during the
 90 image acquisition. This is achieved by using a hardware trigger signal.

¹The stereo-vision system has been calibrated using a chessboard calibration target, following the method described in [9]. The chessboard has been modified to allow automatic initialization.

91 *2.2. Inertial Motion Unit*

92 The IMU is composed of three accelerometers, three gyrometers and three
 93 magnetometers. It can operate up to approximately 120 Hz and gives infor-
 94 mation about the dynamics of our hand-held system :

95 • acceleration $\mathbf{a} = [a_x \ a_y \ a_z]^T$

96 • angular rate $\boldsymbol{\omega} = [\omega_x \ \omega_y \ \omega_z]^T$

97 • surrounding magnetic field direction $\mathbf{b} = [b_x \ b_y \ b_z]^T$

98 The integration of acceleration and angular rate measurements can give
 99 static information – position, attitude – to a certain extent. Indeed, due to
 100 important noise level and biases, position and attitude can drift quickly. It is
 101 yet possible to compose all the dynamics measurements to obtain a relatively
 102 accurate estimation of the attitude of the *IMU* using a Kalman filter [10, 11].

103 Magnetic field is static and consequently its measurement does not drift
 104 incrementally. Accelerometers measure all accelerations applied to them, in-
 105 cluding gravity acceleration. Thus, we are able to express current measure-
 106 ments \mathbf{g}_k and \mathbf{l}_k – respectively current gravity acceleration vector and current
 107 magnetic field direction – with respect to the current estimated attitude.

$$\mathbf{g}_k = R(\mathbf{e}_k) \mathbf{g}_0 \tag{2}$$

$$\mathbf{l}_k = R(\mathbf{e}_k) \mathbf{l}_0 \tag{3}$$

108 where $\mathbf{g}_0 = [0 \ 0 \ g]^T$ is the gravity acceleration vector expressed in world
 109 frame, \mathbf{l}_0 is the measurement of the magnetic field when the inertial sensor

110 is aligned with world frame, \mathbf{e}_k is the current attitude expressed in Euler
 111 angles, and $R(\mathbf{e})$ is the rotation matrix composed from euler angles vector
 112 \mathbf{e} .

113 Angular rate measurements $\omega_{\mathbf{k}}$ are combined at the prediction step of the
 114 Kalman Filter by integration at high rate. Indeed, locally the estimation
 115 does not drift much and the system is considered linear.

$$\dot{\mathbf{e}}_k = \omega_{\mathbf{k}} \quad (4)$$

116 Inertial motion unit frame is noted \mathcal{S} . Transformation between the *coaxial*
 117 *frame* \mathcal{C}_0 and the *inertial frame* is noted \mathbf{H}^{SC_0} :

$$\begin{bmatrix} p^{\mathcal{S}} \\ 1 \end{bmatrix} = \mathbf{H}^{SC_0} \begin{bmatrix} p^{\mathcal{C}_0} \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R}^{SC_0} & \mathbf{t}^{SC_0} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} p^{\mathcal{C}_0} \\ 1 \end{bmatrix} \quad (5)$$

118 with \mathbf{R}^{SC_0} and \mathbf{t}^{SC_0} being, respectively, the rotation matrix and transla-
 119 tion vector of the frame change. This transformation is calibrated by com-
 120 paring attitude measurements and camera poses from the observation of a
 121 known chessboard target. The problem to be solved is then similar to a
 122 Hand-Eye calibration [12].

123 2.3. Light pattern

124 The aim of the system is to allow 3D digitizing in the widest possible
 125 conditions. Consequently we focused on mechanisms to gain independence
 126 to environmental disturbances, lighting being the most important. This is
 127 the reason why the system is equipped with a pattern projector to texture
 128 the scene and facilitate the matching process between images. This projector

129 is a LED matrix coupled with a slide image. This slide is a speckle pattern
130 printed on a glass tile.

131 Moreover, LED rings are added to each camera to be used as secondary
132 light sources in low-lighted environments.

133 Since our system is vision-based, some conditions may still be limiting.
134 Specular reflections will not allow good 3D reconstruction for example. The
135 system tries to adapt its exposure parameters to avoid saturation on images
136 but in the case of intense or large specular reflections, since the system is
137 blinded by light, it will fail to find a good exposure time, and will fail to
138 measure properly. This happens when one scans a specular surface with a
139 frontal point of view : all the light from the projector is reflected towards
140 the system. In this case, the best solution is to scan with a less frontal point
141 of view to direct the reflection away from the cameras. Moreover, since we
142 merge several scans acquired from multiple points of view, a hole created by
143 a specular reflection in an acquisition is filled by the modelling from another
144 acquisition without specular reflection.

145 **3. Operating conditions**

146 In the sequel, strategies and operating modes for 3D modelling using our
147 particular system design are described. An overview is provided and some
148 details about the main methods are given.

149 *3.1. Description*

150 Our digitizing system is an incremental modelling device. The construc-
151 tion of the 3D model is done incrementaly using several raw data acquired
152 sequentially.

153 The aim of our project was to design the system as simply as possible.
154 The user controls the digitizing system using a trigger. A single press on
155 it switches on the device. Prior to doing anything, when starting, the sys-
156 tem launches a sequence for calibration of exposure time and light intensity
157 parameters to fit the environment. This step lasts a couple of seconds.

158 The system is then operational and runs in a preview mode. Three oper-
159 ating modes are available. Figure 2 shows how actions on the trigger drive
160 transitions between each operating mode.

161 [Figure 2 about here.]

162 The starting state is when the digitizing system is off. As we stated above,
163 a short press on the trigger powers the device on to enter a preview state.
164 In this operating mode, raw data are acquired, processed and displayed,
165 and then erased. Pressing the trigger for a longer time allows the actual
166 modelling, that is to say raw data are still acquired and processed but also
167 stored and results of the processing are merged to the results of previous
168 acquisitions.

169 The methods used for incremental modelling were chosen in order to
170 facilitate the overall process. A common difficulty in modelling systems is,
171 for instance, the acquisition of the support side of an object. The object
172 is placed on a table and the support side is not visible unless the object is
173 moved. Most of existing systems require to digitize this side separately and
174 merge it manually during post-processing. The pipeline we propose (Figure
175 3) is designed to allow a fully automatic modelling during acquisition.

176 [Figure 3 about here.]

177 Our methods rely on the use of image processing algorithms and inertial
178 sensing. The main drawback of most methods combining vision and inertial
179 data is that all the processing is based on a fixed global frame. Moving
180 the object is then not possible because it would mean changing the absolute
181 reference. Our pipeline is designed to detect and correct these displacements.

182 For each acquisition time, the digitizing system acquires sequentially, at
183 high rate, two pairs of images – one with the projected pattern, one with no
184 projection – and two attitude estimations synchronised with each pair. A 3D
185 point cloud is created, expressed in the coaxial camera frame. Incremental
186 modelling requires to find the transformation between this camera frame
187 and the global frame of the current model already created. Using attitude
188 measurements and images, a best candidate for alignment process can be
189 quickly chosen. After the alignment step, the 3D points can be merged in
190 a voxel-map-based structure. This is used mainly for visualization, allowing
191 fast and simple sampling of geometric information. Raw data are stored to
192 be used in post-process refinements.

193 If the object has been moved or if the overlap of the current view with
194 the current model is insufficient, no good candidate for registration is found.
195 In this case, a second best candidate search method is used. Of course, if
196 the overlap is too small, this method will fail and the user will be notified.
197 If the object has moved, the inertial measurements become inconsistent with
198 the model frame. To align the inertial frame to the new world frame, a
199 correction transformation related to the object motion has to be estimated.
200 The second method used is based on image indexation to retrieve a good
201 registration candidate from an interest points base. This method is slower

202 than combination of vision and inertial data but is not dependent on a global
203 reference. When a correspondant is found, the registration process gives us
204 the actual transformation between the coaxial camera frame and the current
205 model, allowing us to apply a correction to the attitude estimation from the
206 inertial sensor.

207 When an acquisition is resumed after an interruption, the system checks
208 if the object has moved. Matching is performed using inertial sensing and
209 image detection. If no consistent pose can be found with this method, then
210 the system suggests that the object has moved and the inertial measurement
211 is not usable. The image-only method is then applied.

212 With such a pipeline, the acquisition process is made more user-friendly.
213 Indeed, the modelling can be stopped and resumed easily, and the entire ob-
214 ject can be scanned in the same automatic process by pausing the acquisition,
215 moving the object, and resuming the operations.

216 *3.2. 3D generation*

217 Matching points between stereo images allows an inverse projection to
218 retrieve 3D information from each pair of points. It is the so-called trian-
219 gulation process [13]. When the matching of points or camera parameters
220 are not perfect, optical rays linking camera centers to these points does not
221 intersect. Optimisation or selection heuristic need to be chosen.

222 Considering epipolar geometry, a point in an image and each point on
223 the corresponding epipolar line in the other image are coplanars with optical
224 centers of the cameras. Using stereo rectification [14, 15], matching points
225 lie on the same row in rectified image space, so the intersection between two
226 optical rays always exists. Moreover, epipolar geometry reduces the search

227 space for stereo correspondent to a line of points.

228 Matching is then a two step process : coarse dense pairing, and refine-
229 ment. First step is a global process, trying to match all pixels in an image.
230 The search space in the other image is discretized – to 1 pixel for instance
231 – and is consequently imprecise. This process uses the projected pattern to
232 identify corresponding points without ambiguity.

233 Increasing the search space to be continuous allows the refinement of
234 coarse pairs. Moreover, fine correlation techniques [16] optimize a transfor-
235 mation of the matching pattern to approximate the projective distortion.

236 *3.3. Registration*

237 The online registration process is divided into two main actions :

- 238 • finding the best previous view for a pairwise alignment
- 239 • estimating the transformation for the best fit of both views

240 These methods are only described briefly here and will be discussed more
241 precisely in a forthcoming paper.

242 Finding the best candidate for a pairwise alignment in real time needs,
243 in a way or another, a kind of mapping or classification process. It consists
244 in finding a previously acquired view that maximizes a proximity criterion.

245 In our approach, we use the inertial sensing in first intent to fastly find a
246 candidate in the rotation space. The rotation shift between the two views is
247 then immediatly available. The translation part is found by matching interest
248 points between images. A score is used to decide whether the candidate is
249 good enough or not.

250 It should be noted that 3D data is obtained thanks to the projection of
251 a speckle-pattern onto the scene that is moving with our digitizing system.
252 Matching interest points between images [17] with a moving pattern projec-
253 tion is impossible. This is the reason why each acquisition is composed of a
254 pair of images with projection and a pair without projection. Image points
255 can be matched in images without illumination between two acquisition in-
256 stant n and $n + 1$, allowing to compute pose transformation. It is obvious
257 then that, due to the hand motion, the pose is not exactly the same when
258 acquiring images with projection – from which 3D points are generated –
259 and when acquiring images without projection – from which pose estimation
260 can be found. This introduces an imprecision, acceptable to have a good
261 initial approach, but not accurate enough for an acceptable alignment. A
262 fast registration method – a few iterations of ICP [18, 5] – needs to be run
263 to finish the process.

264 If the search in rotation space gives no result or a result where matching
265 does not pass the score test, then the inertial sensing may have been biased
266 by the motion of the object. The rotation measurement is then forgotten
267 and the candidate is chosen by image indexation, trying to find a previous
268 acquisition whose image interest points match well with those of the current
269 acquisition. This process is not as fast as the first one, which is acceptable
270 for ponctual events like moving the object, but would lead to an important
271 loss in acquisition frequency if used in first intent.

272 In the latter approach, the fast final registration gives an absolute rotation
273 shift between the previous – known – acquisition and the current – biased
274 – one. This shift is used as a corrector for next acquisitions, giving the

275 possibility to use the inertial method again.

276 *3.4. Finalization*

277 The model constructed incrementally in real time is still quite imprecise
278 and can be noisy. To finalize the process, this model needs to be refined and
279 cleaned. This means that the poses and computed 3D points are reestimated
280 together in a kind of bundle adjustment process.

281 The representation of the model is important regarding to the target ap-
282 plication. Meshing is required most of the time. Higher level of representation
283 like surfacing can be needed or other informations like surface texturing.

284 **4. Evaluations**

285 This section aims at showing the potentialities of the scanner within two
286 applications, a qualitative one and a more quantitative one, comparing results
287 to some reference scanning systems of the community.

288 *4.1. Experimental protocol*

289 For our evaluations we scanned several objects with three systems : our
290 hand-held digitizing system and two commercial scanning devices widely used
291 in the community, which will be called *A* and *B* in the sequel. The two
292 commercial devices are fixed scanning devices using optical technology for
293 3D reconstruction. These systems are not clearly identified because, being
294 system designers ourselves, we don't want to bias the purpose of this article
295 to a commercial discussion.

296 In the evaluations we used the Geomagic Qualify v12 software for com-
297 parisons between scans and CAD surfaces. The system *A* does not provide
298 a registration solution so we used Geomagic Studio v12 to align the scans.

299 Scanner *A* is a fixed device, using laser triangulation for 3D measurement.
300 It is an heavy system not easy to handle. System *B* is also a fixed device,
301 lighter than system *A*. It uses fringe projection for 3D measurement and is
302 coupled to a photogrammetry system for localization, requiring that several
303 targets be placed in the scene.

304 Both systems have switchable optics for focal length modification.

305 Operating conditions and specifications of both scanners *A* and *B* are not
306 really comparable to our scanner. However they have been chosen for the
307 announced accuracy and because they are widely used in scientific and indus-
308 trial communities. Technical informations about these devices are provided
309 in table 1.

310 [Table 1 about here.]

311 *4.2. Cylindrical Gauge Block*

312 To get a more defined idea of the relative accuracy of the systems, we
313 started with scanning a cylindrical gauge block used for micrometer calibra-
314 tion. The inner diameter of the cylinder is known (70.004 *mm*) with an
315 accuracy of 1 μm .

316 Evaluation is done using a single acquisition. For each evaluation, the 3D
317 point cloud is registered to a theoretical surface of a 70.004 *mm* diameter

¹Needs to change the focal length to cover all the distances.

²To the *Z* reference plane.

318 cylinder using Geomagic Qualify. Points are projected orthogonally on the
319 surface and the projection distance – the error – is measured. The analyses
320 are based on the mean and standard deviation of the error distances.

321 Table 2 shows the results of this test.

322 [Table 2 about here.]

323 This first test was not intended to make a comparison and ranking of the
324 scanners but to be used as a basis for the evaluation. It tends to give us an
325 absolute reference of the 3D reconstruction process for each device. One can
326 note that our device is in the same accuracy range than system *A* whereas
327 system *B* provides more accurate results.

328 *4.3. Statue*

329 The first evaluation consists in scanning a statue shown in Figure 4.
330 This object is composed of large smooth parts (face, support side, ...) and
331 complex shape areas (cloak mostly).

332 [Figure 4 about here.]

333 At the beginning of the process, the statue is disposed on its support
334 side. We began the acquisition with the face, and then moving to our left.
335 The object has been turned around the vertical axis several times during
336 the process to acquire the cylinder of the head. After a complete turn, we
337 scanned the top of the head. Then the object has been moved to be placed
338 face up. By scanning from the bottom of the face we acquired the support
339 base, which has been completed by placing the object face down and doing
340 the same from the back of the head to the support side.

341 The result of the scan is shown in Figure 5.

342 [Figure 5 about here.]

343 The model is noisy but already recognizable. It still needs to be finalized.
344 We refined the model, cleaned outliers and then meshed. The final result is
345 shown in Figure 6. The texture provided by the images has also been applied
346 to the final model.

347 [Figure 6 about here.]

348 The operation of our device allows us an easy handling which is of great
349 interest in this type of object. The statue has been scanned using the com-
350 mercial scanning devices but handling fixed systems adds difficulties when
351 scanning slightly hidden details like those on the cloak. Figure 7 illustrates
352 a detail of the cloak which suffers information loss in a scan from one of the
353 commercial devices.

354 [Figure 7 about here.]

355 *4.4. Mechanical test piece*

356 For a more quantitative analysis, we used a stamped sheet metal part.
357 To avoid specular effect, the object was mattified.

358 The object is scanned only in its upper face. The sheet being thin, the
359 interest of scanning the opposite face and the sides is questionable considering
360 the resolution and the accuracy of most scanners.

361 [Figure 8 about here.]

362 *4.4.1. Comparison of the three scans with the CAD model*

363 In our test benchmark, we first compared our scan and the scans pro-
364 vided by systems *A* and *B* with the CAD model. Due to mechanical efforts,
365 particularly on release of blank holders, the object has been deformed and it
366 does not correspond anymore to its theoretical CAD model. We registered
367 our test scans to the CAD reference along a same interest zone, favouring
368 the central stamping, supposed to be less deformed. Results are shown in
369 Figure 9.

370 [Figure 9 about here.]

371 The metric used is based on a direct orthogonal distance from scanned
372 points to the reference surface. Figure 10 provides the error distributions for
373 each scan.

374 [Figure 10 about here.]

375 This test allows us to check if there is a bias between several measure-
376 ments of the same object. Comparing the results, our first observation is
377 the similarity of the error maps. With this test, we can see also that the
378 comparison method is stable. Similar data leads to similar registration and
379 projection, considering we used a common alignment reference. This allows
380 us to use this comparison method for the next tests.

381 *4.4.2. Comparison with scan B*

382 The next evaluation focuses on the impact of our finalization step on
383 accuracy. The CAD surface being a bad ground truth we decided to use scan

384 B provided by the more accurate device – according to the manufacturers’
385 specification and the cylinder test – as a reference.

386 [Figure 11 about here.]

387 Figure 11 compares the 3D shape measured with our digitizing system
388 before the post-processing step and after this finalization. Point clouds are
389 projected onto the reference – from system *B* – model.

390 In the raw cloud, a lot of noise appears. This noise is an oscillating phe-
391 nomenon from side to side of the mean surface. Large deviation is observable
392 at the extremities of the model. A boundary effect appears, leading to largely
393 biased measurements in these areas.

394 After the finalization step, noise has been largely reduced, at the point
395 that it is not visible with our colour span. Observing the histogram, we
396 can see that deviation caused by the fast registration method used has been
397 largely reduced by using more accurate registration. The model is more
398 largely comparable to scan B. One can note that boudaries effect, if reduced,
399 still appears.

400 Summary of errors in both steps is given in Table 3.

401 [Table 3 about here.]

402 5. Conclusion

403 In this article we presented our new hand-held 3D scanning device based
404 on vision technologies. We focused the design of this scanner on an easy-to-
405 use scanner. No equipment or markers need to be put in the scene and, being
406 hand-held, the scanner allows a more dexterous manipulation. The operation

407 has been made simple, real-time and mostly automatic. A finalization step
408 allows the creation of a more accurate and usable model.

409 We evaluated the performances of our scanner on several test objects to
410 cover more widely the 3D digitizing applications. We based our evaluations
411 on results obtained by scanning the same objects with well known and widely
412 used commercial devices. Our tests tend to prove that the proposed system
413 matches the requirements of common applications in qualitative terms, accu-
414 racy and usability but with an eased operation allowing to fasten the overall
415 modelling process.

416 Our future works will focus on improving the texturing operation and
417 qualifying more precisely the accuracy and performances of our system. We
418 will also focus our work on trying to reduce noise in the digitizing process.
419 Another improvement can be made by considering ROI-based reconstruction
420 [19]. With such an approach, reconstruction can be processed faster and
421 with reduced noise level due to filtering ill-observed zones outside the central
422 object in images.

423 **Acknowledgments**

424 We thank the IRIT (Institut de Recherche en Informatique de Toulouse)
425 laboratory from Toulouse – France, and the ENIT LGP (Ecole Nationale
426 d’Ingénieurs de Tarbes – Laboratoire Génie de Production) laboratory from
427 Tarbes – France, for the help they provided by scanning our test objects with
428 their commercial scanners (scanners A and B in this paper).

429 **References**

- 430 [1] Pan Q. , Reitmayr G. , Drummond T.W. . Interactive model recon-
431 struction with user guidance. In: Mixed and Augmented Reality, IEEE
432 / ACM International Symposium on. 2009, p. 209–10.
- 433 [2] Curless B. . From range scans to 3D models. SIGGRAPH Computer
434 Graphics 2000;33(4):38–41.
- 435 [3] Matabosch C. , Fofi D. , Salvi J. , Batlle E. . Registration of surfaces
436 minimizing error propagation for a one-shot multi-slit hand-held scan-
437 ner. Pattern recognition 2008;41(6):2055–67.
- 438 [4] Strobl E. , Mair E. , Bodenmüller T. , Kielhöfer S. , Sepp W. , Suppa
439 M. , et al. The self-referenced DLR 3D-Modeler. In: International
440 Conference on Intelligent Robots and Systems. St. Louis, MO, USA;
441 2009, p. 21–8.
- 442 [5] Besl P.J. , McKay N.D. . A method for registration of 3-D shapes.
443 IEEE Transactions on Pattern Analysis and Machine Intelligence
444 1992;14(2):239–56.
- 445 [6] Rusinkiewicz S. , Levoy M. . Efficient variants of the ICP algorithm.
446 In: Third International Conference on 3D Digital Imaging and Modeling
447 (3DIM). 2001,.
- 448 [7] Sandhu R. , Dambreville S. , Tannenbaum A. . Particle filtering for regis-
449 tration of 2D and 3D point sets with stochastic dynamics. In: IEEE Con-
450 ference on Computer Vision and Pattern Recognition (CVPR). 2008,.

- 451 [8] Li H. , Hartley R. . The 3D-3D registration problem revisited. In: IEEE
452 International Conference on Computer Vision. 2007, p. 1–8.
- 453 [9] Zhang Z. . A flexible new technique for camera calibration. IEEE Trans-
454 actions on Pattern Analysis and Machine Intelligence 2000;22:1330–4.
- 455 [10] Marins J.L. , Yun X. , Bachmann E.R. , McGhee R.B. , Zyda M.J. .
456 An Extended Kalman Filter for Quaternion-Based Orientation Estima-
457 tion Using MARG Sensors. In: International Conference on Intelligent
458 Robots and Systems. 2001,.
- 459 [11] Lefferts E.J. , Markley F.L. , Shuster M.D. . Kalman filtering for space-
460 craft attitude estimation. Journal of Guidance, Control and Dynamics
461 1982;5(5):417–29.
- 462 [12] Park F.C. , Martin B.J. . Robot sensor calibration: solving $AX=XB$ on
463 the euclidean group. IEEE Transactions on Robotics and Automation
464 1994;10(5):717–21.
- 465 [13] Hartley R.I. , Sturm P. . Triangulation. Computer Vision and Image
466 Understanding 1997;68(2):146–57.
- 467 [14] Fusiello A. , Trucco E. , Verri A. . A compact algorithm for rectification
468 of stereo pairs. Machine Vision Application 2000;12(1):16–22.
- 469 [15] Bugarin F. , Henrion D. , Sentenac T. , Lasserre J.B. , Orteu J.J. .
470 Optimisation globale polynomiale appliquée à la rectification projective
471 d’images non calibrées (in french). In: Conférence en Reconnaissance
472 des Formes et Intelligence Artificielle. 2010,.

- 473 [16] Garcia D. . Mesure de formes et de champs de déplacements tridimen-
474 sionnels par stéréocorrélation d'images (in french). Ph.D. thesis; Institut
475 National Polytechnique de Toulouse; 2001.
- 476 [17] Bay H. , Tuytelaars T. , Van Gool L. . Surf: Speeded up robust features.
477 In: European Conference on Computer Vision. 2006, p. 404–17.
- 478 [18] Chen Y. , Medioni G. . Object modelling by registration of multiple
479 range images. *Image and Vision Computing* 1992;10(3):145–55.
- 480 [19] Li W. , Schütze R. , Böehler M. , Boochs F. , Marzani F.S. , Voisin Y.
481 . Preprocessing of region of interest localization based on local surface
482 curvature analysis for three-dimensional reconstruction with multireso-
483 lution. *Optical Engineering* 2009;48(6).

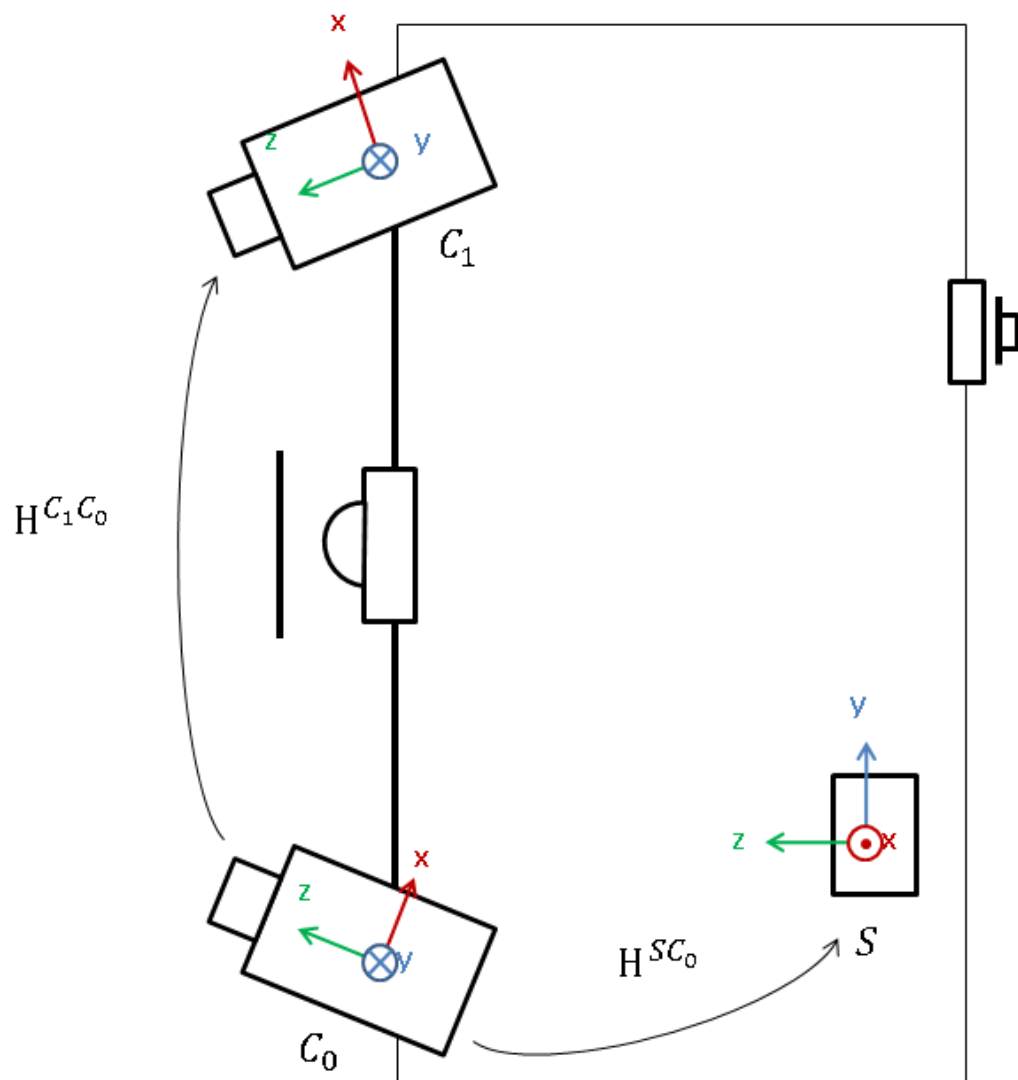


Figure 1: Design of the system with sensor frames and transformations.

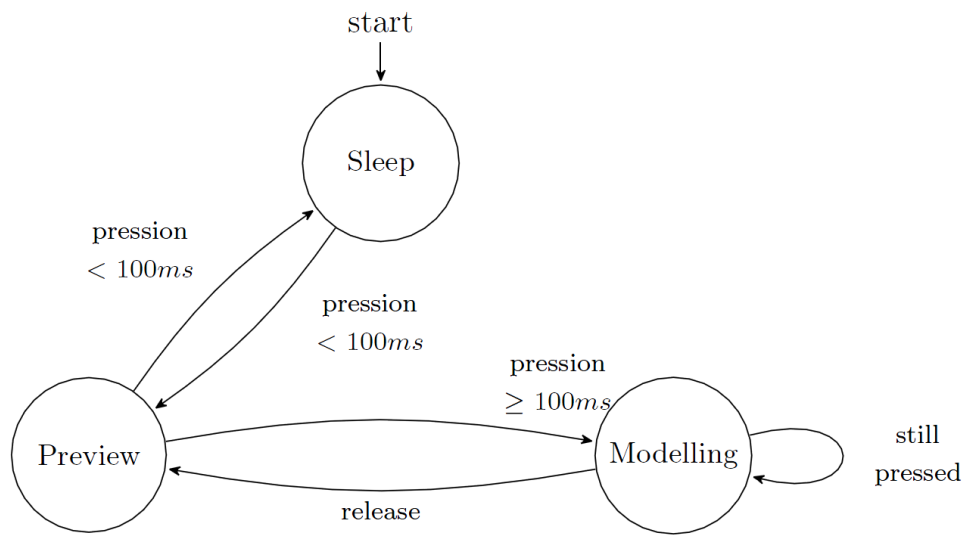


Figure 2: State diagram of trigger control

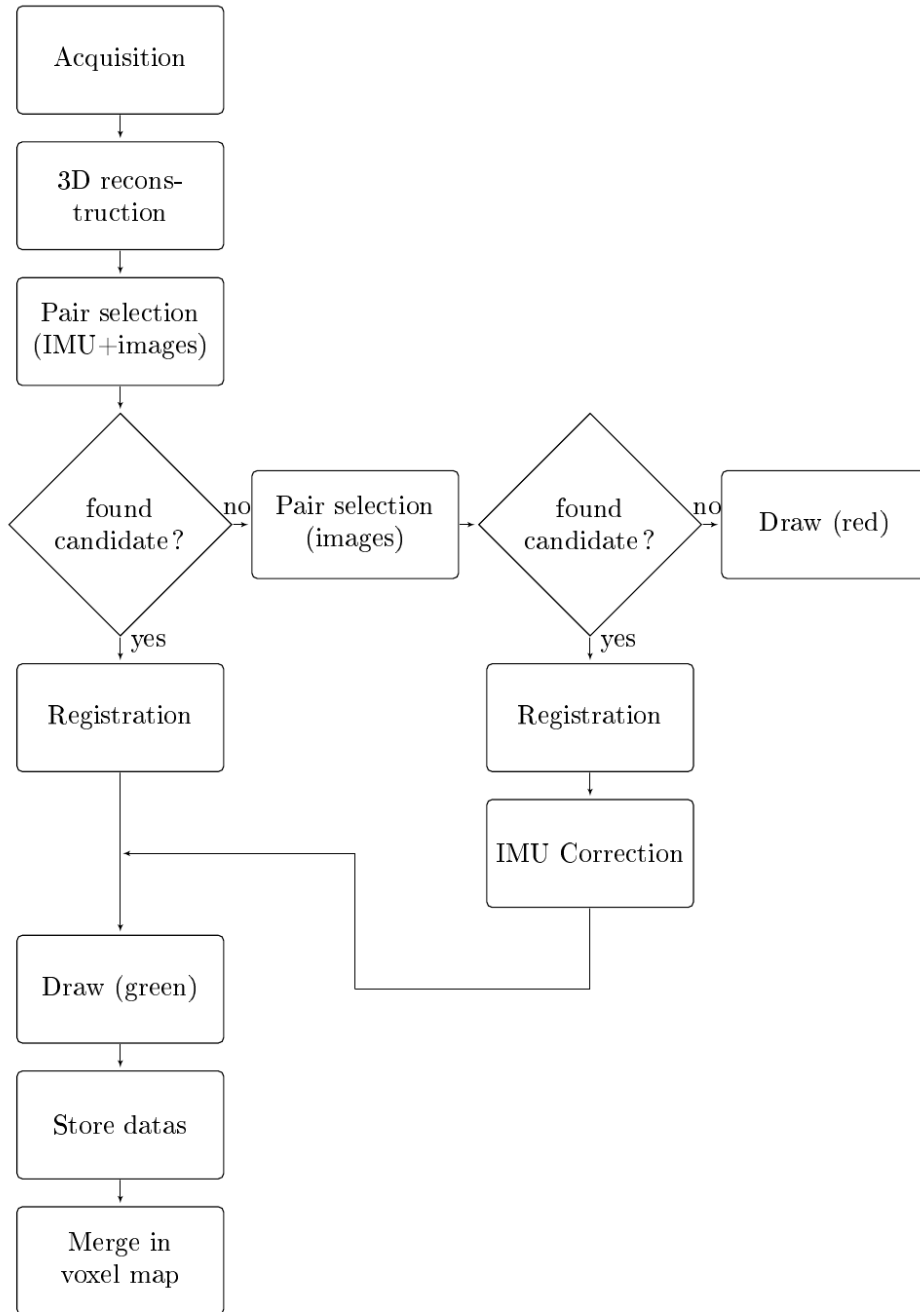


Figure 3: Operation diagram for single acquisition processing.



Figure 4: Test statue.

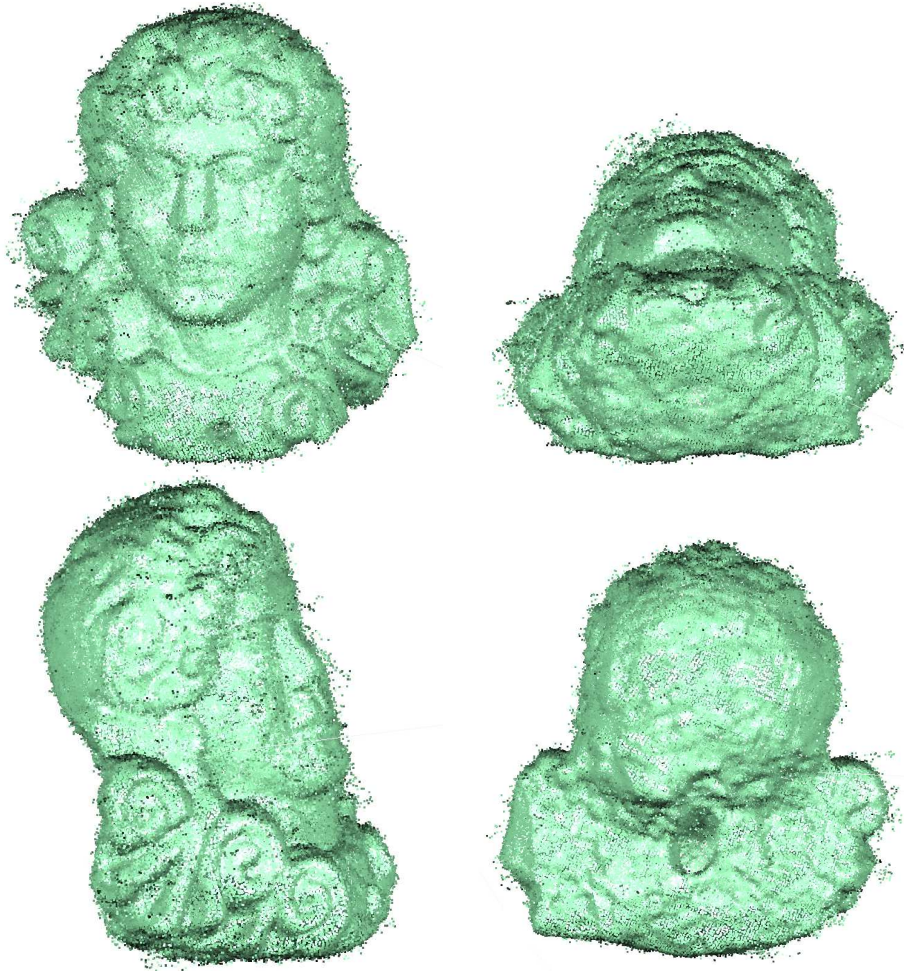


Figure 5: Result of the scan without finalization step.



Figure 6: Finalized and textured scan of the statue using our system.

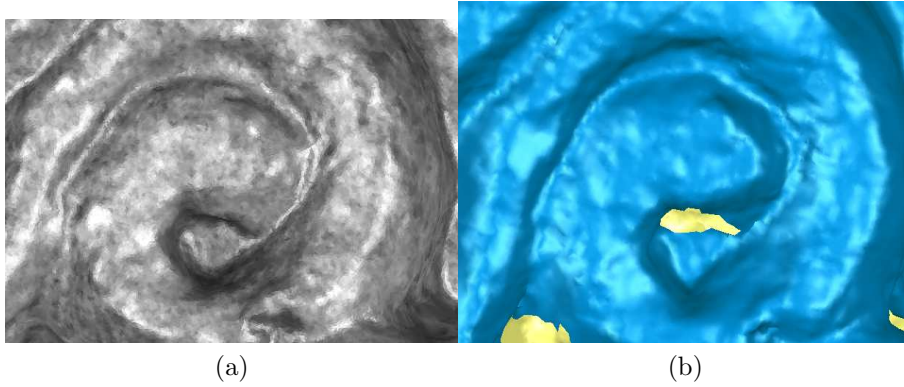


Figure 7: Detail of the object. In complex areas of the object, some faces can be hard to observe due to occlusions. (a) Observing complex surfaces with our system is eased by its hand-held operation, allowing a more dexterous manipulation. (b) Scan from system A, with a less dexterous operation information, losses can occur.

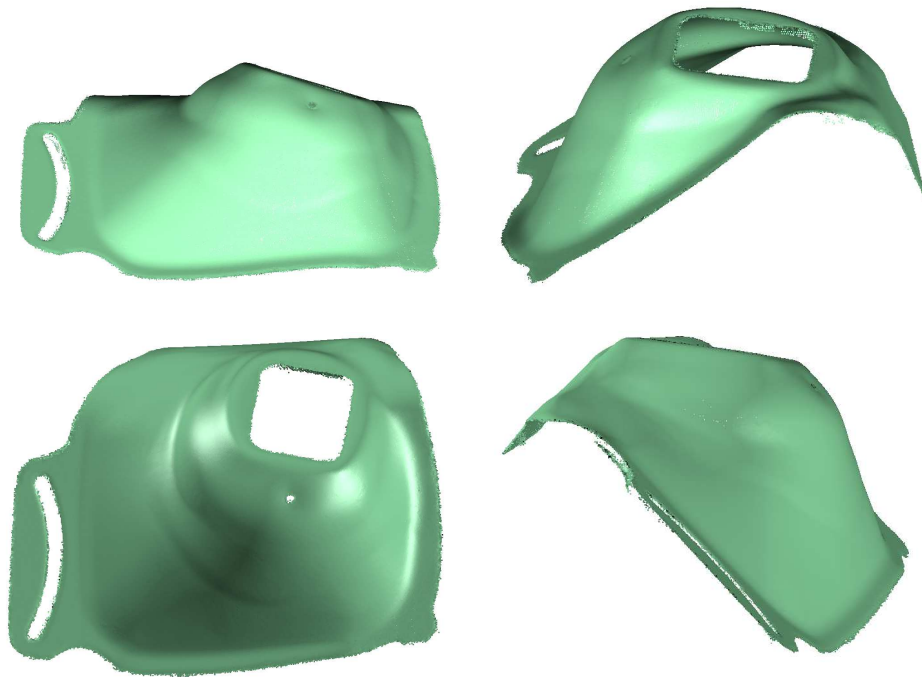


Figure 8: Stamped sheet metal part used for our tests and resulting point cloud provided by our system.

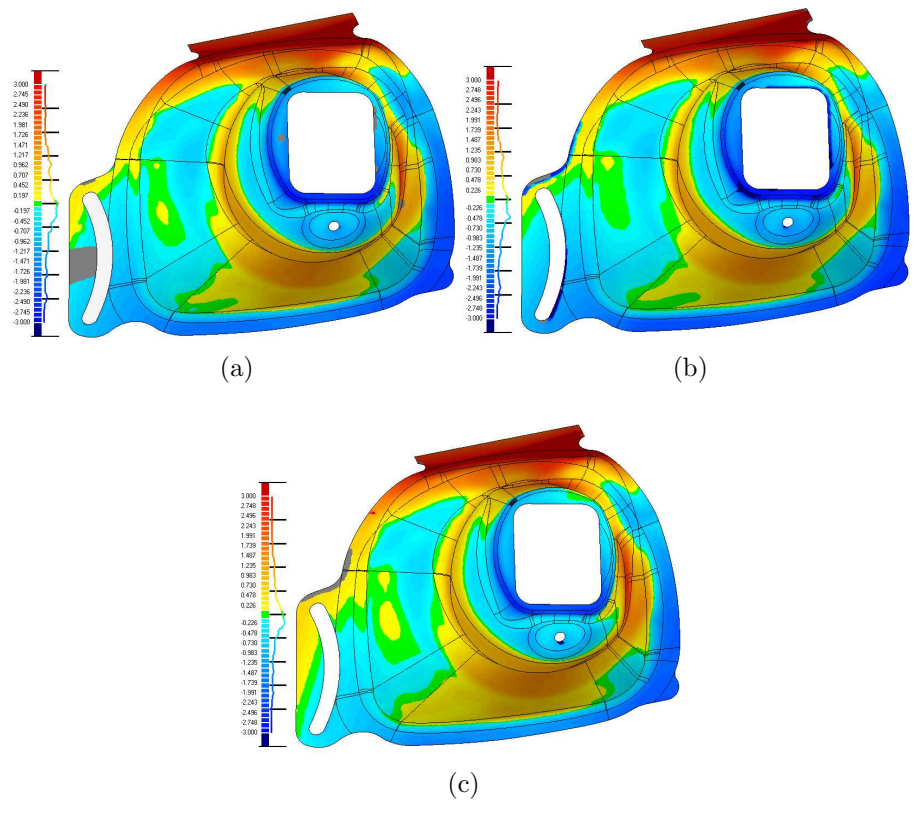


Figure 9: Comparison of scans with theoretical surface : (a) Scan using system A ; (b) Scan using system B ; (c) Scan using our hand-held scanner.

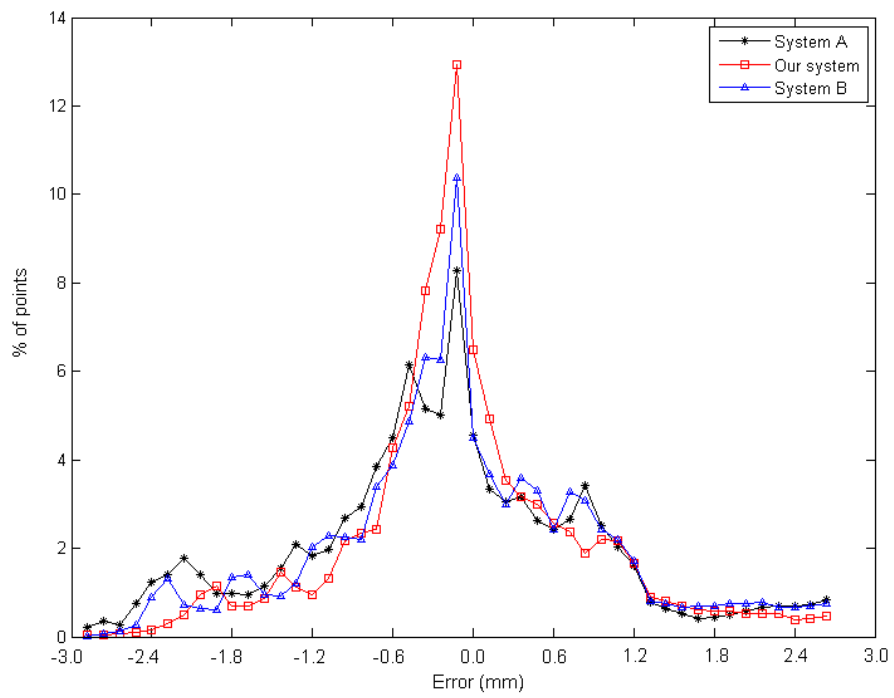
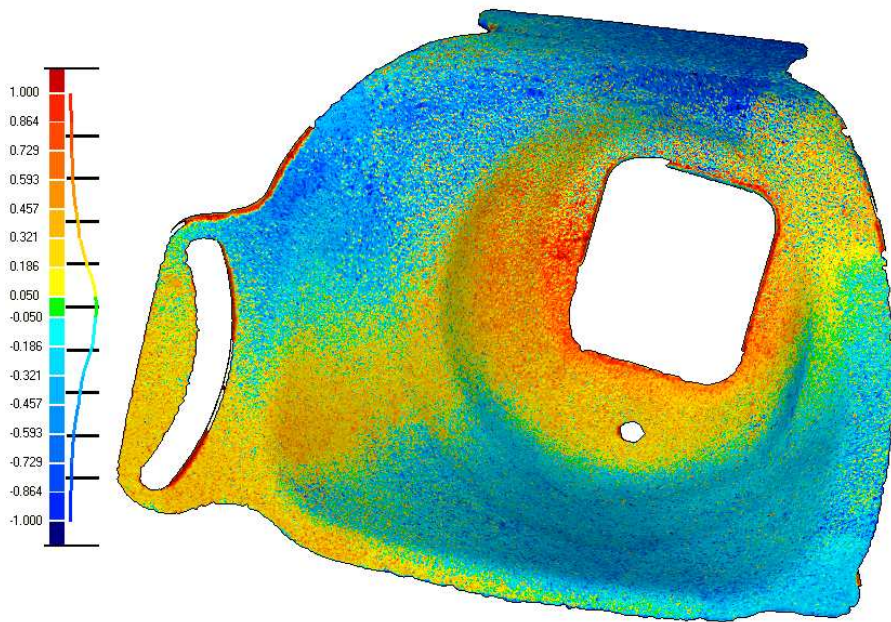
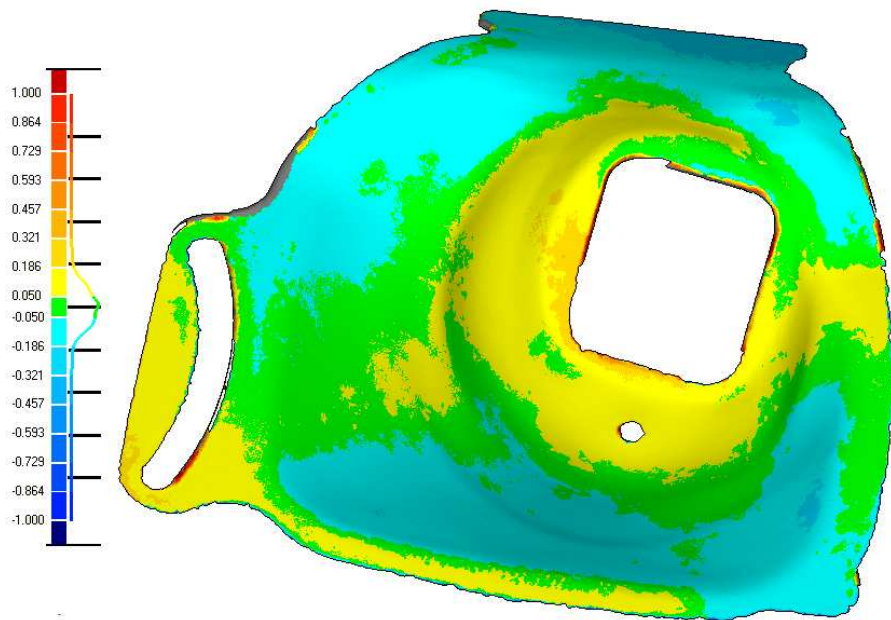


Figure 10: Distribution of error between 3D scans and the CAD surface.



(a)



(b)

Figure 11: Comparison between our scan and scan B : (a) before finalization ; (b) after finalization.

	A	B	Our system
Motion	Fixed	Fixed	Hand-held
Technique	Laser triangulation	Structured light	Structured light
Registration	Manual	Photogrammetry	Automatic
Weight (kg)	15	7.4	1.8
Dimensions (mm)	$221 \times 412 \times 282$	$490 \times 300 \times 170$	$220 \times 240 \times 90$
Focal length (mm)	14	16	8
Operating distance ¹ (m)	0.6 – 1.2	0.7 – 2	0.35 – 0.5
Measurement volume (mm)	from $111 \times 83 \times 40$ to $1196 \times 897 \times 800$	from $135 \times 108 \times 108$ to $1700 \times 1360 \times 1360$	from 10cm^3 to 1m^3 NC
Accuracy ($\mathbf{X}, \mathbf{Y}, \mathbf{Z}$) ² (mm)	$\pm (0.22, 0.16, 0.10)$	NC	NC
Point spacing (mm)	NC	0.08 – 1.0	0.3
Measuring noise (mm)	NC	0.002 – 0.02	0.1

Table 1: Systems specifications (Commercial systems specifications are provided according to manufacturers’ data sheets).

	A	B	Our system
Mean (mm)	-0.00115	0.00003	-0.00115
Std. dev. (mm)	0.02764	0.00508	0.02480
Max. error (mm)	0.10238	0.02162	0.08649
Min. error (mm)	-0.09151	-0.02569	-0.09392

Table 2: Evaluation results scanning a known gauge block with the three systems.

	Before finalization	After finalization
Mean error (mm)	-0.004	-0.001
Std. dev. (mm)	0.373	0.125

Table 3: Errors between our scan and scan B