

# Building and Tracking Root Shapes

J.J. Jacq, C. Schwartz, V. Burdin, R. Gérard, O. Rémy-Néris, C. Lefèvre, C. Roux

**Abstract**—An algorithm aiming at robust and simultaneous registrations of a sequence of 3D shapes was recently presented in *IEEE T-BME 55(5)* by Jacq *et al.* Along the way, this algorithm has to carry out an implicit representation of their common root shape. A particular emphasis was put on the median consensus shape, a specific type of root shape. Unlike this previous work, mainly focussing on the algorithm foundations while dealing with very specific applications examples, this paper attempts to show the versatility of the root shape concept through a set a three problems involving a wider scope of application. The first problem copes with the design of prosthetic cortical plates for the hip joint. It shows how an explicit reconstruction of the root shape, coming with its consensus map, could bring out an intermediary anatomical support from which could be made pragmatic choices performing a tradeoff between morphological, surgical, and production considerations. The second problem addresses in vivo real-time shoulder biomechanics through a miniature 3D-video-camera. This new protocol implicitly operates through root shape tracking of the content of virtual spotlights. It is shown that the current medical-oriented protocol, while operating within expert offices through low-cost equipments, could challenge high-end professional equipments despite some limitations of the 3D-video-cameras currently available. The last problem deals with respiratory motions. This is an auxiliary measurement required by some medical imaging systems that can be handled as a basic application case of the former new protocol.

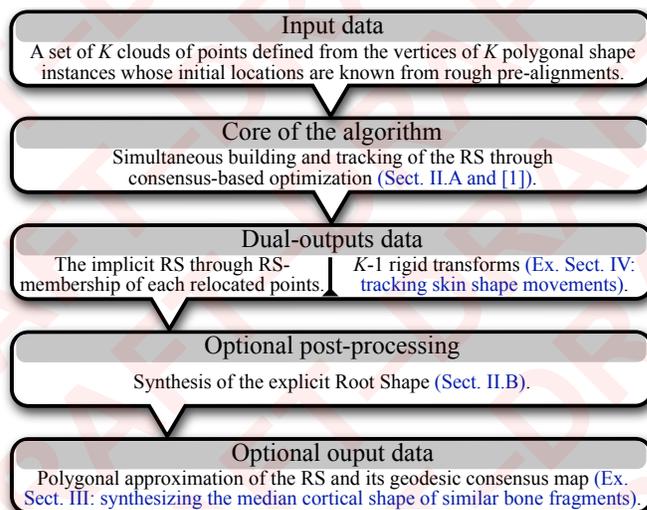
**Index Terms**—Prosthetic plate, markerless tracking, 3D-video-camera, shoulder kinematics, respiratory tracking, 4D medical imaging, root shape, simultaneous registrations.

## I. INTRODUCTION

**R**OBUST matching of multiple observations of a noisy, artifacted, and unevenly truncated 3D shape is a challenging task that may be involved in various biomedical engineering problems. A survey of the main algorithmic strategies can be found in [1]. While addressing individual movement measurements of carpus and tarsus bones observed in vivo through standard 3D medical imaging techniques, a new iterative algorithm operating robust simultaneous registrations through a median consensus rule was presented in [1]. The algorithm was thus initially devised w.r.t. a very specific problem. However, by introducing two novel medical applications, both of

larger scope, this paper attempts to show that this approach can go way beyond the former problem. For semantic generalization purposes, the median consensus rule is now defined as a specific setup of the consensus rule. Likewise, the Median Consensus Shape (MCS) becomes a specific instantiation of its underlying Root Shape (RS) generic concept. Thus, the algorithm is now referred to as the Root Shape Tracking (RST) algorithm.

The aim of this methodological-oriented paper is to promote the significance of the RS concept, a.w.a. the effectiveness of the RST algorithm. Fig. 1 summarizes the data-flow involved by the proposed framework. It emphasizes the dual nature of the simultaneously extracted though an RS-based processing. In one hand, it outputs (1) relative locations of the competing instances w.r.t. their common RS. On the other hand, it outputs (2) an implicit description of the RS. In other words, RS-building tasks and RS-tracking tasks are the two faces of a same problem. The inclusions of two innovative biomedical protocols taking separately advantage of the dual ability of the processing framework are thus made complementary w.r.t. to the main objective of this work. The first protocol addresses bone prosthetic design whereas the second addresses in vivo kinematics from skin shape instances.



**Fig. 1.** Overview of the general data-flow involved by an imaging protocol built on the RS concept. The parentheses summarize its relationships with the content of the paper.

The paper is organized as follows – see also Fig. 1. Section II.A summarizes the main lines of the Root Shape Tracking algorithm previously described in [1] and introduces some additional working options. Section II.B describes the new post-processing step addressing the explicit reconstruction of the root shape. Section III describes an RS-building-based protocol. Its application involves the design of a prosthetic

Manuscript submitted October 11, 2008, accepted April 8, 2009.

J.J. Jacq (corresponding author, [JJ.Jacq@telecom-bretagne.eu](mailto:JJ.Jacq@telecom-bretagne.eu)), V. Burdin and C. Roux, are with Institut TELECOM, TELECOM Bretagne, Technopôle Brest-Iroise, CS 83818, 29238 Brest Cedex 3, France.

R. Gérard and C. Lefèvre are with Service de Chirurgie Orthopédique, Traumatologique et Réparatrice, CHU La Cavale blanche, Bd Tanguy Prigent, 29200 Brest.

O. Rémy-Néris is with Service de Rééducation Fonctionnelle, CHU Morvan, 2 av Foch, 29200 Brest, France.

C. Schwartz, and previous authors, are with Institut National de la Santé et de la Recherche Médicale, INSERM, U650, LaTIM, CHU, I3S, 5 av Foch, 29609 Brest cedex, France.

cortical plate through the explicit reconstruction of the root shape of the great trochanter. Section IV describes an RS-tracking-based protocol addressing in vivo markerless kinematics through measurements issued from a miniature 3d-video-camera; its applications mainly focus on scapula kinematics a.w.a. respiratory gating. Section V concludes with some perspectives.

## II. ROOT SHAPE TRACKING AND BUILDING

### A. Overview of the RST algorithm

The RST algorithm [1] is primarily a point matching process operating simultaneously on  $K$  geometric observations. The input data consists of set of  $K$  clouds  $\{P_k, k = 1, \dots, K\}$ , with points  $P_k \equiv \{p_{k,i}, i = 1, \dots, N_k\}$ . Each cloud stands for the geometric component of a triangulated shape whose network description – that is, its edges list and its triangles list – is implicitly known. Thus, the RST algorithm now always operates in polyhedral mode. A new optional input may specify a priori knowledge through intrinsic normalized weights  $\{W_{0k}, k = 1, \dots, K\}$  with  $W_{0k} \equiv \{w_{0k,i} > 0, i = 1, \dots, N_k\}$ . Internally, this static weight is simply managed as the multiplicity level of a point. The RST optimization process introduces a transversal information flow that attempts to manage a single  $K$ -dimensional M-estimator in an experimental way. Along the way, the optimization process has to carry out an implicit representation of the common Root Shape of the  $K$  instances through a consensus-based merging rule. Let  $\alpha \in (0,1]$  denote the normalized rank of the consensus level requirement. In [1], while setting  $\alpha = 0.5$ , a particular emphasis was thus put on the Median Consensus Shape as a specific type of Root Shape. The outputs of the algorithm are (1) the RS-membership maps  $\{\hat{W}_k, k = 1, \dots, K\}$  with  $\hat{W}_k \equiv \{\hat{w}_{k,i} \leq w_{0k,i}, i = 1, \dots, N_k\}$  and (2) the rigid pose estimations  $\{\hat{T}_k, k = 1, \dots, K\}$  of the  $K$  initial cloud locations, both w.r.t. their common Root Shape. The coordinate system embedding the Root Shape is kept aligned with the one of the first instance. The inverse of the successive rigid transformations received by this instance are thus reapplied to the  $K$  instances so that  $\hat{T}_1 = I$  holds continually. When  $K = 2$ , the RST algorithm degenerates to a robust and symmetric pairwise matching algorithm called Robust Reciprocal Matching (RRM) afterwards.

The main parameter is the outlier error scale  $\lambda$ . Previous parameters  $\lambda'$ ,  $It_{\max}$  are respectively defaulted to  $\lambda$ , 1 and are no longer under user control. A new parameter  $\delta$  controls an implicit fuzzy extension of the shape borders along tangent planes while transiently enabling point-to-plane metric w.r.t. any target point located on a shape border: the initial target is then shifted to the closest location defined by this metric. The part of the actual shift length in excess w.r.t.  $\delta$  is managed as a usual matching errors and thus defines an additional normalized weight multiplied to the current weight of the border target point. The minimal error gain  $\mu$  (typ. 0.1%) still controls convergence w.r.t. each instance modification. However, unlike the applications described in [1], the current number of instances may now become large ( $K > 1000$ ). Thus, the deterministic global convergence criterion has to be replaced by

a statistic criterion related to the inter-quartile  $IQ$  of the current distribution of the  $K$  error gains. Convergence occurs if  $IQ < 2\mu$  with any error gain modulus larger than  $2IQ$ .

### B. Synthesis of the explicit RS

An optional post-processing step enables an explicit reconstruction of the RS. In essence, the approach follows early works (e.g., [2], [3]) addressing volumetric-based fusion of ranges data. The explicit RS is built as the zero-set of the signed-distance function computed over a regular grid whose domain encloses the whole set of registered instances. A downward-octree-oriented hierarchical scanning enables to disregard subdivisions of the grid-cells that do not intersect the RS and thus saves much of the computational overhead. The final construction step outputs an explicit polyhedral description of the zero-set through a standard iso-surface extraction techniques. Thus, w.r.t. the RST framework, the remaining specific problem comes down to perform an accurate estimation of the signed distance w.r.t. each corner location of an incoming test-cube. This local task is formally similar to the RST step that collects the target points associated to a source observation point. However, as the latter is now *external* w.r.t. the intervening data points, the RST merging rule becomes irrelevant. The centroid of the  $K$  targets, a.w.a. its robust normal giving access to the sign of the distance, is thus computed through a new merging operator. The latter sorts the projections of each target location over their respective grid axis and gives them an individual cell wide defined by their final RS-membership  $\hat{w}$ . The merging weights are then estimated through a polished version of the Tukey's trimean L-estimator. For each axis, this involves the geometric construction of two asymmetric ramps, each starting from opposite sides of the location of the  $Q_2$  quartile with elevation 1.0, over the segment enclosing the concatenated  $K$  cells. A typical setting of the ramps would then involve elevations 0.5 on both  $Q_1$  and  $Q_3$  locations. However, the current merging rule sets the two extrema elevations to 0. Each cell trapezoidal area, excluding negative elevations, then defines the per-axis point merging weight. The mean of these three per-axis weights provides the actual merging weight of the point. The local *consensus level*  $\Psi$  of the RS is defined as the sum of these  $K$  merging weights. A modified Marching-Cubes algorithm delivers the surface consensus-level  $\Psi$  as an additional channel linked to the vertex-list. To reduce parallax effects while computing the distance grid, a robust *condensation* of the  $K$  registered clouds can be done right from the end of the last RST iteration. This attenuation of the Gaussian noise is simply done by moving, independently, each source point toward its virtual RST target point along a path fraction proportional to its RS-membership  $\hat{w}$ . A former reconstruction of the MCS of the tarsus was presented in Fig. 13.c–d of [1] p. 1631.

## III. SEEDING THE SHAPE OF A PROSTHETIC PLATE

### A. Medical objective

While performing total hip arthroplasty, on purpose of redistributing the forces loads over the remaining parts of the proximal bone structures of the femur, the adjunction of an

auxiliary cortical prosthesis may be required. The main anatomical structures of the regions under consideration involve areas of muscular insertion – e.g., the lesser and greater trochanter a.w.a. the linea aspera. To estimate the morphological constraints linked to the outline of such an auxiliary prosthesis – here modeled as a rigid cortical plate – thirty Caucasian femoral dry bones have been digitalized and then segmented to recover the 3D shapes of their outer cortical interfaces.

### B. Challenging aspects of the problem

A standard approach would proceed through the computation of the mean shape while then performing a Principal Components Analysis (PCA) of the dispersion of the samples w.r.t. this mean shape. If, as usual, the protocol operates through geometric morphometry [4], [5], the geometry of the input shapes has to be described through a compact set of landmarks. Thus, as the selection criterion of a marker location is mainly defined in concern with its retrieval accuracy, the resulting set of locations – typ. local bumps linked to tendon insertions – may not truly account for global perturbations of the shape. Indeed, depending on muscles powerfulness, areas of muscular insertion may be subject to important deformations and this has long been recognized as the source of the noticeable inter-individual variability of the proximal segment of the femur [6]. Marker locations, most of which being linked to these local anatomical shapes, are therefore imprinted by this random saliency. This makes them expressing more local, instead of global, information.

Whether an accurate PCA can be carried out, a panel of plates fitting the main perturbations of the mean shape would then have to be defined. However, while considering production cost – i.e., attempting to limit available plate instances so as to primarily fit global scale a.w.a. left-right asymmetry – no options are left to perform a fine-grained coverage of these main modes – which, from the previous discussion, can be expected to be numerous. Thus, this rules out possibilities of performing a full account for the main modes of the perturbations through a very limited set of plates.

### C. Method and results

It is expected that working through the root shape and its consensus map together, a more *pragmatic* design process may then be carried out while taking into account anatomical, surgical, and economical requirements. More precisely, the proposed strategy guesses that, while performing through a few consensus-map-based perturbations of the root shape, one may achieve an operational trade-off between these conflicting requirements – i.e., coping with a high inter-individual dispersion while offering a subset of prosthesis at reasonable production costs.

For normalization purpose, the femoral length is defined as the length of a segment joining up the more distal point of the intercondylar fossa and the more proximal point of the greater trochanter. The thirty bones are then uniformly scaled to exhibit a common Unitary femoral Length (UL). Surface parts whose normalized axial abscissa along the reference segment is greater than 0.24 UL are discarded. To focus on the regions that are in concern with the prosthetic functionality, a semi-

interactive coarse preselection is then carried out. This mandatory preprocessing step aims at delineating the greater trochanter along with its top inner side – including the trochanteric fossa – a.w.a. the proximal metaphysis while excluding the lesser trochanter. The resulting morphological unit can then be described as a “spoon”. A selection example is depicted in Fig. 2.a.

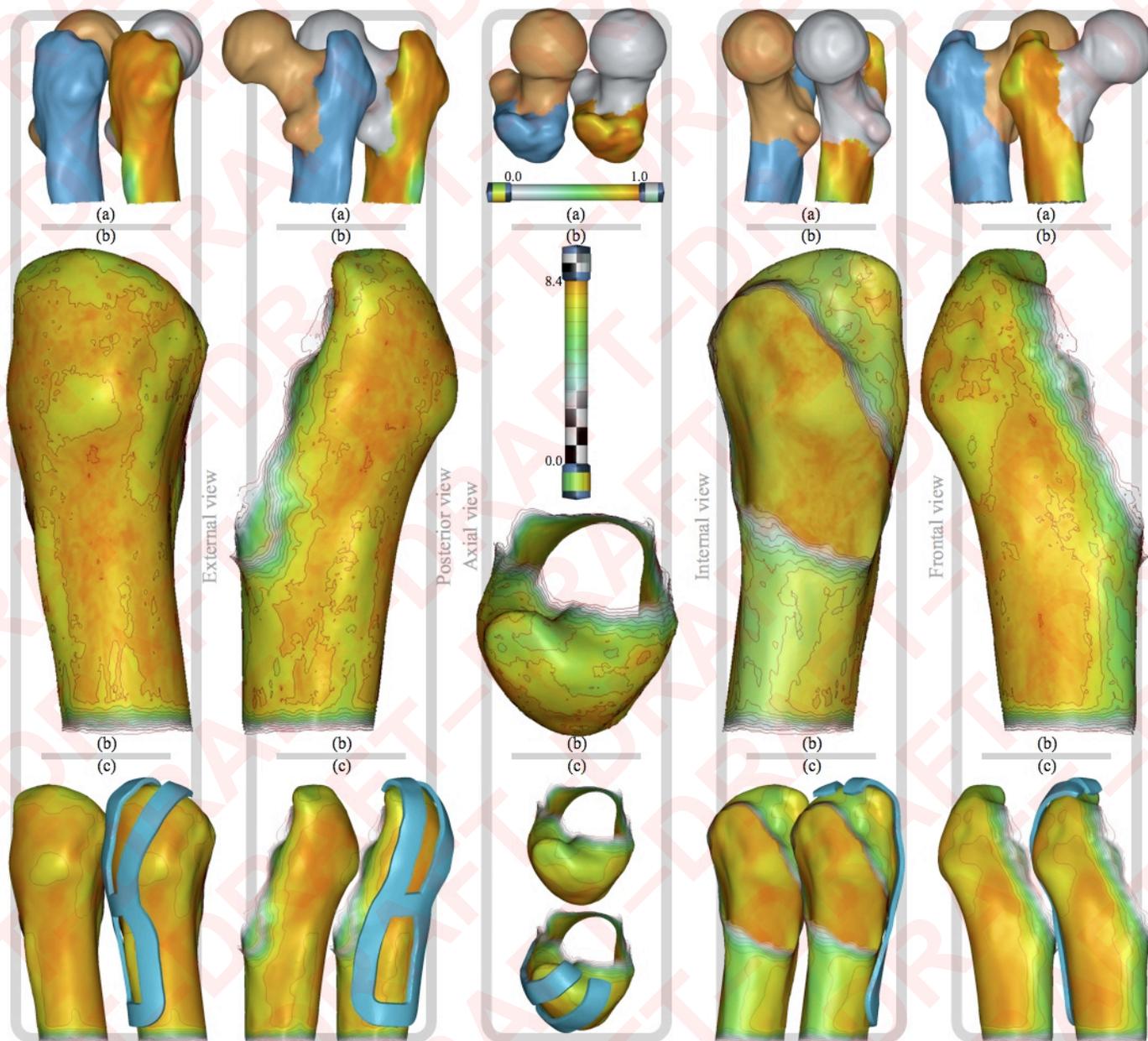
The simultaneous tracking and building of the root shape operates from the initial guess provided through alignment of the principal axis of the thirty preselections. The consensus-based RST algorithm is tuned to follow the median rule (i.e.,  $\alpha = 0.5$ ) with main parameters  $w_0 \equiv 1$   $\lambda = 5$  and  $\delta = \sim 0.01$  UL. The resulting MCS is depicted in Fig. 2.b. The geodesic field mapped over the MCS accounts for its local consensus level  $\Psi$ . The step of the MCS reconstruction grid was  $1.5 \cdot 10^{-3}$  UL. As expected, Fig. 2.b shows that the regions of poorest consensus focus on main insertion areas – including the linea aspera. As the geometric transforms encode a rigid model, the algorithm logically chooses to dismiss the internal side of the shaft so as to focus on the “spoon” regions.

Fig. 2.c depicts the same MCS with a geodesic-filtered instance of its raw consensus level. This filtered consensus level  $\tilde{\Psi} = A^{r_2}(O^{r_1}(C^{r_1}(\Psi)))$ , is obtained while cascading three geodesic operators. First, an opening of geodesic radius  $r_1$  is applied to  $\Psi$ . Thereafter, the resulting field undergoes a closing step over the same geodesic radius. Finally, a diffusion is applied while averaging the previous field over a geodesic radius  $r_2$ . The setting involved in Fig. 2.c are  $r_1 = 5 \cdot 10^{-3}$  UL and  $r_2 = r_1 / 2$ . The synthetic shape (blue object) embodies a morphological ground truth example. Its path selection criterion focusses on stability w.r.t. a specific osteological objective. This synthetic model is issued as a molding expanded through normal-oriented offsets modulated by the local level of the filtered consensus. This seeding shape can then be seen as the very first step of the design of the prosthetic plate. The next steps – including the setting of a meaningful algebraic relationship between  $\tilde{\Psi}$  and the offset – must consider additional and transversal aspects, involving both surgical-oriented and production-oriented requirements, that are not in the scope of this paper.

## IV. TOWARD MARKERLESS ANALYSIS OF SKIN MOTIONS

### A. Objective and strategy

Nowadays, due to availability of powerful stereophotogrammetric systems, technics aiming at performing in vivo, real-time, and accurate measurements of human movements requires prior affixation of skin markers. However, apart from both time-consuming-related and poor reproducibility shortcomings of its affixing step, the main drawback of this approach is related to Soft Tissue Artifacts (STA). Leardini *et al.* [7] discuss this problem and perform a survey of the solutions proposed to cope with STA. Thus, it is well known that some biomechanics measurements are distrusted by these artifacts. As a matter of fact, while analyzing a few set of static poses observed through MRI modality, Matsui *et al.* [8] studied sliding discordancies between successive locations of a spherical skin marker and the underlying scapula and reported errors

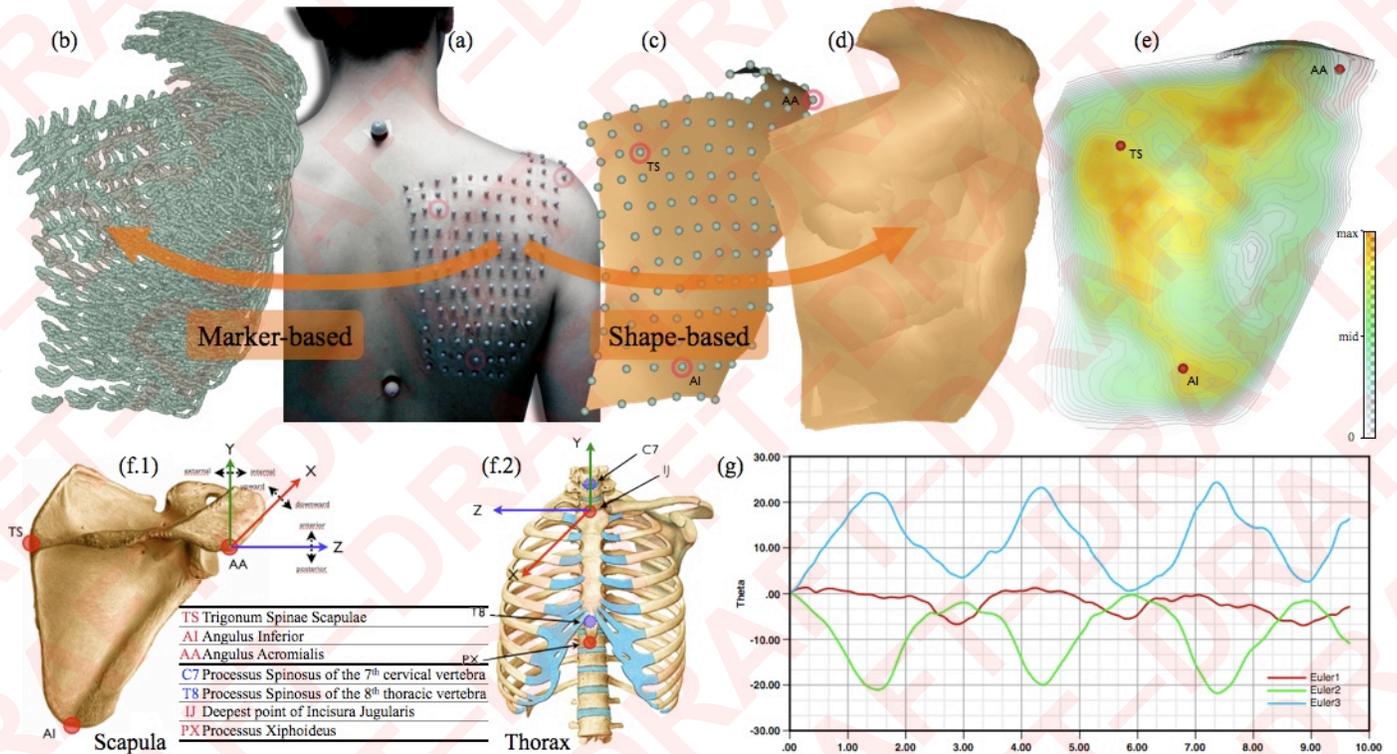


**Fig. 2.** Synthesis, with application, of the root shape of thirty great trochanters through the median consensus rule. Each cartouche-like column of this 3x5 matrix depicts, through orthographic projections, the shapes seen from one of five main anatomical views : *external*, *posterior*, *internal*, *frontal* and *axial*. The top row (a) displays one of the input shapes with its resulting MCS's membership map – *i.e.*, a color map accounting for a probability level ranging from 0 to 1. This input shape emerges, a posteriori, as the more representative w.r.t. the common MCS. The blue parts outline the result of a semi-interactive preprocessing step aiming at preselection of its regions of interest. The middle row (b) displays the MCS of the corresponding morphological unit. In addition to the color, the consensus level also controls the transparency level of the MCS – the maximum of the consensus level amounts to 8.4. The bottom row (c) depicts the same MCS with a geodesic-filtered instance of its raw consensus level along with a synthetic shape (blue) embodying the morphological ground truth w.r.t. the seeding of the shape of a prosthetic plate.

ranging up to 39 mm w.r.t. the acromion palpating locations.

From the methodological point of view, the protocol discussed in this section is a first attempt to develop a new approach enabling to get rid of these usual skin markers and their STA-related artifacts while challenging the performance-cost ratio of available optoelectronic systems. Its main idea proposes to track of the 3D skin shape over a whole – or part of it – body segment while superseding usual physical markers by surface-related patches. As the former are usually punctual – *i.e.*, a collection of 0D geometric objects – the latter can be seen as 2D virtual markers. Therefore, although it deals with a more specific objective, this approach follows the same Mark-

erless Motion Capture (MMC) methodological class that [9]. To cope with STA, the current working hypothesis then guesses that some parts of the complex shape fluctuations may convey an external fuzzy imprint of the rigid bone sliding under the skin. Thus, performing a consensus-based global analysis of the 4D (3D+T) data in order to extract the root shape should filter out elastic components of the shape movements. Instantaneous locations of the body segment are then implicitly recovered from the tracking of the Frame of Reference (FoR) of the root shape. Before extracting motions measurements, FoR trajectories are low-pass-filtered through a 6DoF-constrained Butterworth. The rotation measurements



**Fig. 3.** Estimating scapulothoracic movements through an optoelectronic tracking device. This figure depicts the processing steps of three flexion-extension periods by the IMCP algorithm. The data acquisitions convey up to 1200 instances dispatched over 10 sec. Unlike markers-based tracking (b), an interpolating surface (c) is built for each instance locations. Upon simultaneous registrations (e) of the set of surfaces (d), the colors mapped on the resulting skin's MCS (e) account for its consensus level – white means zero consensus (*i.e.*, unreliable parts) whereas red means maxima of the consensus. Views (b) and (d) depicts the whole sweep in thorax coordinate system. The common MCS (e) is drawn at rest location ( $t = 0$ ). View (f) summarizes the definitions of the skeletal landmarks of the thorax and the scapula w.r.t. to the International Society of Biomechanics (ISB) recommendations. While looking back at (e) it appears that only two – of three – scapula ISB landmarks are found to be in good accordance with the consensus map levels. View (g) depicts an YXZ-order Euler's factorization of the rotations of the scapula w.r.t. the thoracic-FoR (*Euler1*: Retraction/Protraction; *Euler2*: Lateral/Medial rotation; *Euler3*: Anterior/Posterior tilt).

will attempt to follow the recommendations of the International Society of Biomechanics (ISB) [10] (see also Fig. 3.f). However, as the investigated protocols attempt to be markerless and do not provide omnidirectional views, the measurement conventions can not follow these markers-related recommendations strictly. This should not be a major impediment since current measurements focus on rotations only<sup>1</sup>. Furthermore, unlike the previous application case (section III), the explicit reconstruction of the root shape can now be left as an optional byproduct.

The main obstacle currently preventing scrutinization of this prospective approach has to do with the lack of 3D sensors performing both dense spatial sampling and dense temporal sampling – *i.e.*, a Dynamic 3D Shape Tracking Device (DSTD). As specified in [11], such a *template* device denotes a class of markerless systems able to undergo video sampling rate, at least 60 Hz, while still performing dense, regular, and simultaneous sampling of an evolving 3D shape with a spatial accuracy compatible with common medical requirements, that is, in the range of a few millimeters.

The application involved by sections IV.B–E addresses an open problem: *in vivo* tracking of scapula movements through skin shape motions. Subsection IV.B shortly summarizes a previous attempt at solving this problem with the RST-algorithm whereas subsections IV.C–E describes an innovative protocol putting forward a new technology, 3D-video-cameras, which is on the way to challenge the DSTD speci-

cations. In these experiments, a 3D-video-camera is setup to provide both coarse temporal sampling (20 Hz) and dense spatial sampling (5 mm). The section IV.C brings out technical information and then presents the results. The section IV.D discusses about current limitations of this new protocol whereas section IV.E summarizes its promising results *a.w.a.* its rational evolutions. To assess the versatility of this new protocol, the subsection IV.F addresses another skin-based problem, still mainly related to the thorax: respiratory gating.

#### B. Shape-based tracking emulation through an optoelectronic system

A comprehensive description of the corresponding experiments can be found in [11]. Starting from rest position, the patient performs forward flexions in the sagittal plane. The region of the skin surface masking scapula locations is evenly sampled ( $\sim 15$  mm edge length) through a set of 123 (4 mm) markers. Successive locations of these markers are measured with a VICON optoelectronic system. The latter comprises seven infrared video cameras MX13 (1.3 Mpix.) operating at 120 Hz. A triangulation of the 123 vertices provides a polyhedral approximation of the corresponding patch. Inner (*resp.* border) vertex locations are then randomly shifted along tangent planes (*resp.* border edges) to de-correlate successive instances of the moving patch. The results are depicted in Fig. 3. The corresponding RST parameters was  $w_0 \equiv 1$ ,  $\alpha = 0.5$ ,  $\lambda = 5$ ,  $\delta = 15$  mm.

However, some comparisons with palpation-based measurements, show that the RST results still underestimate the amplitude of scapula motions. The RST algorithm does not provide a decisive improvement w.r.t. its initial guess, that is, the principal axes of the patches. Indeed, the spatial sampling could not be made dense enough to fully recover the skin fluctuations. Therefore, despite the fuzzy extension of the borders of the patch ( $\delta = 15$  mm), the perimetric-related matching errors still get influence over those of the inner parts where insightful shape fluctuations are expected to take place. Furthermore, a lateral physical extension of the patches may have performed better but was not fully tractable through the configuration of the optoelectronic system.

### C. Shape-based tracking using a 3D-video-camera

Upon discarding sensors unable to simultaneously sample all parts of their view field, widespread non-contact 3D sensors providing dense sampling can be classified mostly into two technologies: active structured-light systems and multi-view (typ. stereo) systems. As an example, the visual hull of the body can be reconstructed, with 10 mm resolution, from simultaneous multi-views provided by 8 colors VGA cameras operating at 75Hz [9]. In order to create a robust real-time 3D scanning system, Weise *et al.* [12] recently propose an up-to-date hybrid system combining both structured-light and stereographic approaches through a stereo phase-shift method. The system operates at 17 frames per second (fps) – most parts of the real-time reconstruction algorithm being implemented on the GPU – and performs reconstructions with a 0.125 mm accuracy at 1.1 m distance. However, the whole system simultaneously requires to dispatch one video projector, two synchronized monochrome cameras, and one color camera while still providing a limited view-field – the latter spreads over  $\pm 0.3$  m away from the center of the view-frustum.

Conversely, as an emerging technology, 3D-video-camera appears to be of special concern w.r.t. the concretization of the generic 3D real-time device introduced above and in [11]. Such a camera embeds its own light source and performs the direct physical measurement of the time of flight  $\tau$  of the source light backscattered by the target surface. Thus, along with the usual pixel-related information (location vector, sight line unit vector, intensity or RGB channels), each pixel now delivers a depth measurement  $d = c\tau/2$  where  $c$  is the speed of light. Moreover, this technology enables to pack miniature 3D-video-cameras with an integration level similar to usual laptop 2D-cameras. There are two camera types.

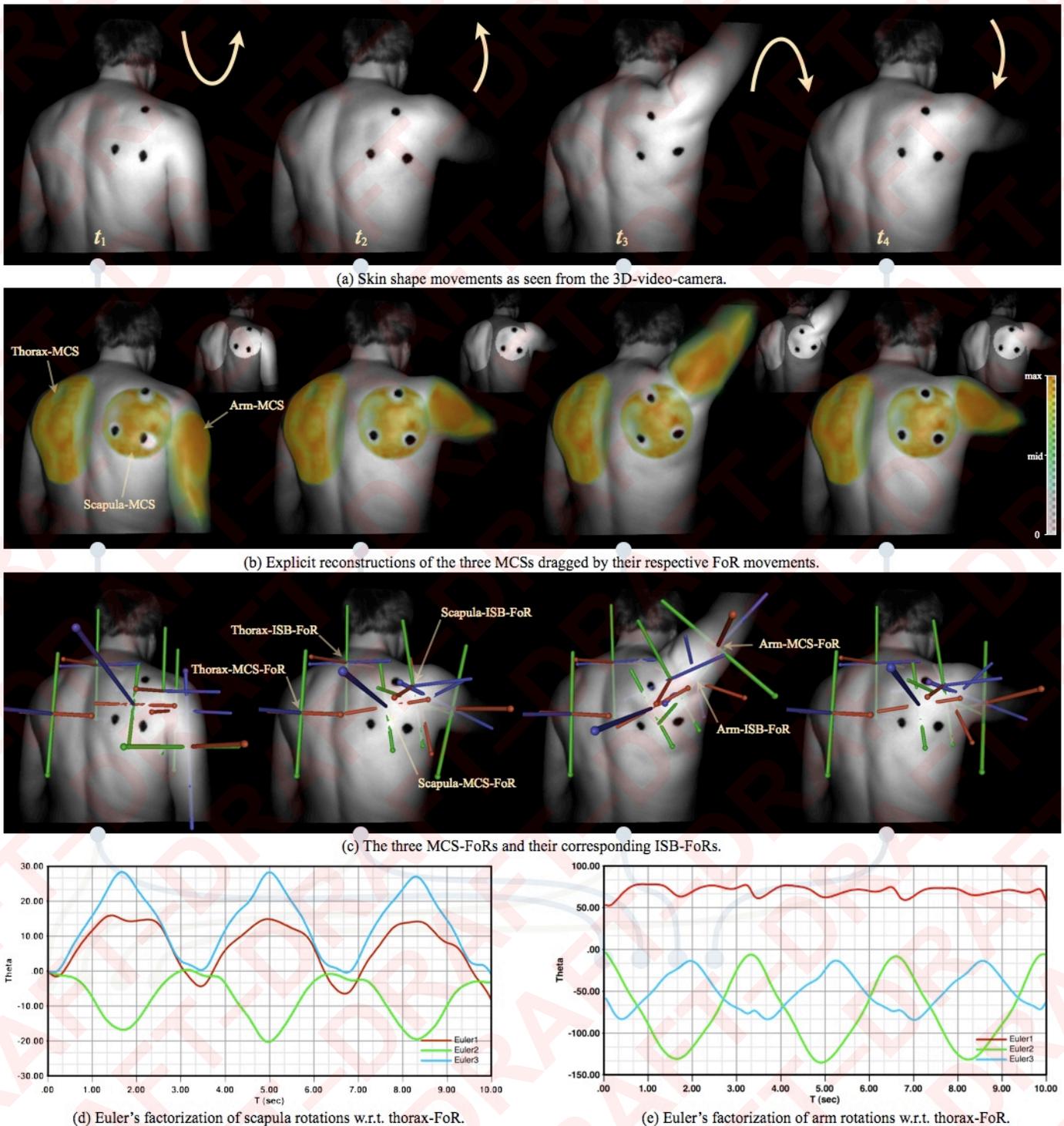
The first type, known as Z-camera [13], recurrently sends light pulses. Its technology has most to do with the radar fundamental concept and may thus be called scanner-less lidar. To maximize the accuracy of depth measurements, pulses must be generated through very-high-speed optical shutter to minimize the duration of the transition edges. The depth accuracy of the camera does not directly depend on the distance from the camera. The setting of the depth measurement window width determines the resulting depth resolution. Current specifications of the ZCam<sup>TM</sup> miniature model are: up to 60 fps, maximum depth window 0.5 to 5 m with 1–2 cm depth precision (which may come down to a few millimeters whenever the 8-

bits-encoded depth window may be stretched accordingly), binocular 2D-3D (i.e., including both a QVGA RGB-sensor and a depth-sensor). Despite these high performances, due to its limited depth precision, current applications of this model should be more in concern with 3D-TV and PC-video-gaming.

The second type, known as ToF-camera [14], continuously performs a modulated light emission. The simplest working configuration (i.e., mono-camera) operates through sinusoidal modulation (frequency  $\nu$ ) of the source amplitude. A real-time correlation of the returned signal w.r.t. the signal generated by the local oscillator carries out an estimation of the phase shift  $\varphi$ . This phase shift enables to compute the corresponding travel time  $\tau = (2\pi\nu)^{-1}\varphi$ . Unlike the first camera type, this technique brings out a major limitation: to perform unambiguous measurements, the view-field must not contain reflecting objects whose depth exceeds  $d_{\max} = c/2\nu$ . However, typical values of  $\nu$  – e.g., 20 MHz ( $d_{\max} = 7.5$  m) or 80 MHz ( $d_{\max} = 1.857$  m) – involve depth limitations that are still compatible w.r.t. medical applications.

By design, 3D-video-cameras suffer from problems inherent to radar systems – e.g., multiple ray trajectories may end on the same pixel – and may then lead to complex and convoluted artifacts. Moreover, the actual depth resolution may be poor w.r.t. transversal resolution and the SNR of the depth measurements decreases with the object albedo. The experiments described below make use of an up-to-date miniature ToF-camera<sup>2</sup> with  $d_{\max}$  set to 7.5 m. As a matter of fact, with this model running through optimal conditions at 20 fps, the standard deviation of the depth noise on central pixels amounts to 6 mm for objects distant from 1 m and gradually increases with distance. Moreover, in common situations, additional environmental-dependent noise sources may be involved [15]. These shortcomings made the requirement for the robustness of the RST approach even more suitable.

A 3D-video-camera-based experiment was thus setup to mimic the previous shoulder experiment. Fig. 4 summarizes the protocol a.w.a. the measurement results. A video giving a comprehensive overview with additional information is also available from [16]. Upon correct placement of the camera, its virtual sight line is horizontal and goes through the TS scapula landmark – the latter being distant from 0.8 m w.r.t. the origin of the camera-FoR. This line is rotated by  $\pi/3$  w.r.t. to the sagittal plane so as to focus on the right shoulder. The remaining freedom degree of the camera is defined by aligning its Y axis with the vertical axis of the room. As the corresponding view-field includes both the proximal and medial parts of the arms, the protocol will also attempt to simultaneously perform the tracking of the arm to provide extra measurements without additional constraints. The frame data comes out as 3D points cloud sets  $\{p_i, \phi_i, i = 1, \dots, N_0\}$ , with  $N_0 = 2534$  points-per-frame, where  $p$  is a point location within the camera-FoR and where  $\phi$  is an additional scalar channel providing the intensity of the returned light signal. The implicit connectivity provided by the 2D grid sensor makes obtention Delaunay's triangulation of the clouds a trivial task. From its relative placement, upon discarding points whose depth is greater than 1.5 m, the camera provides meshes with 5 mm median edge length



**Fig. 4.** Estimating scapulothoracic and thoracohumeral movements through a 3D-video-camera device. This figure, focussing on three flexion-extension periods, depicts the processing steps of a Root-Shape-based protocol. The data acquisitions convey up to 200 instances dispatched over 10 sec. View (a) depicts four samples of the 3D video sequence - they are linked to the four key positions of the first cycle. The optical axis of the 3D-video-camera is setup through  $33^\circ$  of external rotation w.r.t. to the vertical axis - the mean skin distance along this axis is 0.8 m. The three circular footprints – made of black makeup – enable to compute an initial guess for the scapula FoR. View (b) depicts the MCS-based tracking of skin shape areas. These areas are defined through intersections with virtual spots (left: thorax; middle: scapula; right: arm). The background view display the virtual spots. These three spots are built through a cascaded analysis; first, the thorax spot is drawn from the camera-FoR, in the second step, the scapula spot is drawn from the thorax-MCS-FoR, then, in the last step, an hemispherical spot focusing on the arm is drawn from the scapula-MCS-FoR. The MCSs are depicted through the conventions previously described by Fig. 3. View (c) displays the three MCS-FoR and their kinematically equivalent ISB-FoR companions. However, as the thorax face is unknown, the thorax-ISB-FoR can not strictly adhere to the ISB recommendations (see Fig. 3.f.2) : its origin is translated from landmark IJ to C7. Nevertheless, as the movement estimations currently focus on rotations, this FoR shift has no incidence. View (d) depicts an YXZ-order Euler's factorization of the rotations of the scapula w.r.t. the thoracic-FoR (*Euler1*: Retraction/Protraction; *Euler2*: Lateral/Medial rotation; *Euler3*: Anterior/Posterior tilt). View (e) depicts an YXY-order Euler's factorization of the rotations of the arm w.r.t. the thoracic-FoR (*Euler1*: elevation plane,  $0^\circ$  is abduction,  $90^\circ$  is forward flexion; *Euler2*: Elevation (negative); *Euler3*: axial rotation, exo- or external-rotation and endo- or internal-rotation). An Mpeg4 video describing each processing step throughout the full 3D-video sequence can be downloaded from [16]. Red, green, and blue tubes depict X-, Y-, and Z-axis respectively through pins-like conventions so that the head of an axis points to the negative direction.

– this is also the transversal resolution within the skin area of the scapula.

The tracking algorithm must take care of the view- and time-dependent 3D-silhouettes. These are ribbon-like artifacts inherent to range measurements. Even though these artifacts always imply very low signal intensities, whether very low intensities account for such artifacts is still context-dependent. An intrinsic normalized weight is thus pre-computed to encode how well a vertex is likely to pertain to ribbon-free areas. Let  $\Lambda$  denote the mean of the normalized quality<sup>3</sup> of the triangles sharing vertex  $p$ . The raw weight  $\tilde{\phi} = \Lambda^2 \phi$  of vertex  $p$  is then defined as its intensity measurement  $\phi$  weighted by  $\Lambda^2$ . Let  $w = (\tilde{\phi} - \tilde{\phi}_{\min}) / (\tilde{\phi}_{\max} - \tilde{\phi}_{\min})$  denote the normalized initial weight where  $\tilde{\phi}_{\min}$  and  $\tilde{\phi}_{\max}$  denote the extrema of this raw weight w.r.t. the vertices of the current instance. Let  $Q_1^w$  (resp.  $Q_2^w$ ) denote the first (resp. second) quartile of  $w$ . With reference to the corresponding input of the Root Shape Tracking algorithm, the intrinsic weight  $w_0$  of the vertex  $p$  is then defined as  $w_0(w) = (1 + \tanh(k(w - w_r))) / 2$  with  $w_r = Q_1^w - \beta_0(Q_2^w - Q_1^w)$ . The parameter  $k$  – i.e., the slope at location  $w = w_r$  – is adaptively defined through an implicit benchmark. The adaptive parameter  $k$  is thus defined as the smaller slope such that both  $w_0(Q_1^w) \geq 0.8$  and  $w_0(Q_2^w) \geq 0.99$  requirements hold simultaneously. The sequence depicted in Fig. 4 was processed with  $\beta_0 = 1$ . The RST algorithm is tuned to follow the median consensus rule (i.e.,  $\alpha = 0.5$ ). Its main parameters are defined with  $\lambda = 3$ ,  $\delta = 15$  mm. Initial guesses of the thorax-FoR and arm-FoR locations are both provided by principal axes – or by the RRM algorithm for disk-like shapes linked to the scapula-FoR.

Like the previous protocol, the new protocol (Figures 4.a–c) addresses scapulothoracic rotations. However, it also provides extra-measurements focussing on rotations of the thoracohumeral pseudo-joint. It operates through hierarchical definitions of Virtual Spotlights (VS). First, a static ellipsoidal thorax-VS is drawn for the camera-FoR. Upon matched through the RST algorithm, its MCS-FoR becomes available. Then, a circular scapula-VS is drawn from the thorax-MCS-FoR. The scapula-MCS-FoR is then provided by the RST algorithm. Lastly, a hemispherical arm-VS is drawn from the scapula-MCS-FoR and the arm-MCS-FoR is then computed. The three black cosmetic disks are auxiliary scapula markers that are discussed in section III.D. The RST algorithm operates from the sequence of skin-patches intersecting a VS. To reduce the RST computational cost, the triangle meshes – a.w.a. their vertex scalar channels  $\phi$  and  $w_0$  – may be simplified before entering the matching process. With respect to measurements of the

scapula rotations, comparisons between Fig. 3.g and Fig. 4.d show that this new protocol might challenge optoelectronic-based protocols. Unlike previous experiments, a significant retraction-protraction component (Euler 1) is even detected. However, amplitudes are still underestimated. Indeed, scrutinizing relative locations of the medial makeup-based circular mark in the companion video [16] makes this noticeable without resorting to palpations of the scapula. Regarding rotations of the arm (Fig. 4.e), even if these are secondary aspects of the current protocol, it was still able to recover both elevation (Euler 2) and elevation plane (Euler 1) angles<sup>4</sup>. However, as the remaining  $\pm\pi/3$  oscillations are mainly artifacts, it fails w.r.t. axial rotations of the arm (Euler 3). Therefore, these results seem promising but still express current device limitations that are discussed below.

#### D. Discussion

For didactic purposes, the shapes (Fig. 4.a–c) were previously described from the actual 3D camera location through the same perspective angle with a texture defined by the camera intensity channel. In turn, Fig. 5 now depicts two of these views through a morphometrical-oriented setup – i.e., orthographic projections with intrinsic OpenGL shading – along a view axis orthogonal to the shoulders plane. In spite of the realtime filtering steps applied to the raw depth measurements (a 3x3-XY median spatial filter a.w.a. a low-pass temporal filter, both provided by the camera), both a high level noise and artifacts still remain. So, having recourse to the root shape tracking approach appears very sound. Thus, tracking the thorax-FoR from an initial guess issued from the camera-FoR becomes an easy RST task.

Unlike the experiments described in section III, the transversal resolution is now adequate. However, because of the substantial noise level spoiling depth measurements, tracking the skin shape fluctuations over the scapula area from an initial guess issued from the thorax-FoR still remains a very difficult objective. Indeed, a fully markerless protocol could not be implemented. The proposed scapula-adapted protocol requires two steps. To get better initial guesses of the scapula-FoR locations, the first step makes use of three disk-like (17 mm diameter) black marks imprinted on the skin. The second step only considers the parts of the scapula-VS not obscured by these three markers and is thus similar to the RST processing step of the other body segments. To preserve shape information over the more informative skin regions, these markers are stamped away from usual landmark locations. The black color makes it easy to segment these three markers from the intensity channel. Conversely, as the corresponding echo signal vanishes when the target reflectivity decreases, the distance measurements become much more unreliable there. Thus, making use of the RST algorithm is still a requirement while addressing this preliminary first step.

In connection with the movement artifacts previously underlined, the arm sequence shows another major difficulty. For each pixel, the 3D-video-camera under examination requires four consecutive echo samples to estimate the phase shift between the source ray and its echo. While operating at 20 Hz, 12.5 ms are then devoted to each sample acquisition – of

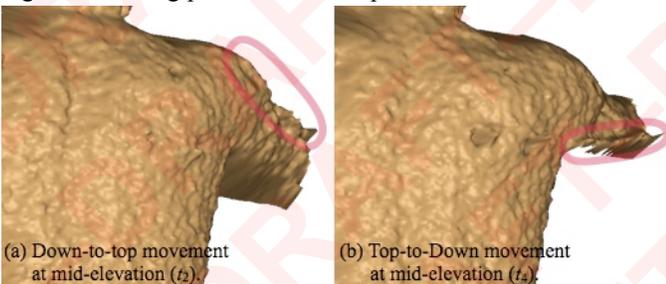


Fig. 5. Overview of the available geometric information linked to sample states  $t_2$  et  $t_4$  (Fig. 4). The red curves underline the main movement artifacts.

which no more than 8.2 ms can be made available to the actual integration time. This makes the depth measurement not fully instantaneous w.r.t. to medial and distal parts of the moving arm where abrupt changes may occur within a frame period – especially near its outline borders. Moreover, the global 3D shape of the resulting artifacts may appear mutually coherent along half a cycle. As this may challenge the median consensus rule, related bias effects are not fully cancelled and their residuals give rise to some hysteresis-like phase shifts of the rotation components around the arm axis (see Euler 3, Fig. 4.e). Whenever more attention is paid the arm-MCS-FoR movements, the video [16] makes this phenomenon more obvious. An alternative protocol including the skin shape of the elbow should be less sensitive to these problems. Some perspectives also addressing these problems are discussed below.

### E. Perspectives

In the near future, biomechanics measurements issued from 3D-video-camera should be able to achieve performances gained by professional optoelectronic systems. Indeed, unlike the optoelectronic tracking protocol, the 3D-video-camera-based protocol is even able to recover some additional rotation components. As discussed above, the main drawbacks of the 3D-video-cameras currently available refer to high-speed movement artifacts and significant noise level. This has prevented these first experiments to successfully match the DSTD requirements that would support a *fully* markerless protocol. Meanwhile, the current protocol seems very promising and still offer significant advances from many sides:

- it may challenge, even upgrade, results issued from up-to-date specialized materials;
- the protocol is noninvasive – it involves neither contacts nor manipulations of the patient;
- it involves very low cost materials w.r.t. current specialized systems;
- it provides a fully ambulatory desktop miniature system able to operate in the physician's office;
- as the 3D-video-camera works with its own near-infrared source light, it may tolerate, to some extent, moderate room lights.

Moreover, two major evolutions may help to shortly relax parts of the current drawbacks. The first evolution has to do with incoming new 3D-video-camera models that will perform depth measurements with better accuracies. The second evolution immediately comes out from visual comparisons between intensity-based-shading renderings (Fig. 4) and OpenGL-based virtual-shading of the same surface (Fig. 5). This makes obvious that bridging between depth and intensities measurements would greatly improve the SNR and filter out major movement-related artifacts. As a matter of facts, according to principles of the Shape-from-Shading (SfS) theory, Böhme *et al.* [17] recently carried out remarkable improvements while de-noising static ToF measurements. However, as this de-noising step is time consuming, addressing time sequence still requires more efficient SfS-based algorithms. Conversely, due to its intrinsic 4D-level robustness, the root shape approach should cope with suboptimal SfS-based filtering results a.w.a. simplistic assumptions w.r.t. the albedo model.

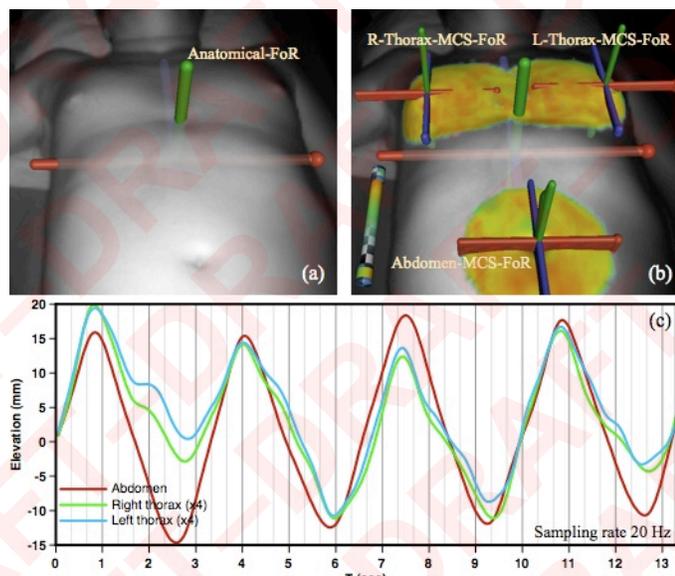
While tracking some limb segments, a measurement protocol operating through a sole ToF-camera still provides a limited redundancy level. As shown previously w.r.t. to the axial component of the arm rotations, this protocol may be prone to deliver inaccurate results. Obviously, using multiple 3D-video-cameras simultaneously should overcome this problem. However, while operating through sinusoidal modulation of the light sources, concurrent operation of several ToF-cameras would lead to incorrect phase shift estimations. Büttgen *et al.* [18] propose to employ pseudonoise-modulated intensity signal to minimize inaccuracies resulting from multi-camera interferences and are then able to manage measurements with up to five ToF-cameras operating simultaneously in laboratory environment. Meanwhile, their measurements become much more sensitive to ambient noise – e.g., background light.

### F. Tracking Respiratory Motions

Respiratory motions may be involved as an auxiliary measurement required by common medical imaging modalities – e.g., SPECT, PET. In this context, the patient is laid down at rest position on a table. A protocol addressing real-time respiratory gating through ToF measurements has been recently proposed by Schaller *et al.* [19]. The authors firstly determine the horizontal background plane before examination then, during examination, remove background plane data and, finally, divide the camera view field in two axial-oriented half parts, that is, the thorax and the abdomen. Two horizontal planes, whose elevations versus time determine the two gating signals, are then robustly fitted to their respective remaining data points – the latter being weighted by their own light intensity. Measurements was found to correlate well with those provided by an external respiratory gating system.

Some component of the biomechanics-oriented protocol described in the previous section may obviously offer an alternative to the former pattern-matching-based approach. Indeed, the thorax tracking strategy may be applied to respiratory objectives without any change. As the body is now static in the camera-FoR and respiratory motions are mainly translations along the vertical axis, the tracking task becomes much more easy. Moreover, as the noise and the resolution are both stationary, processing successive elevation estimations through the RRM algorithm, that is, performing an independent localization of each instance versus reference instance  $t_0$ , nearly achieves the accuracy of the RST algorithm as far as spot contents are not unevenly clipped. Real time computation of a spot elevation signal is thus in the scope of current processors.

The tests are depicted in Fig. 6. As the camera field of view is too small ( $40^\circ$  along Y-axis) to be fitted into the tunnel of common medical imaging device while still covering more than 50–60 cm along any patient trunk, its line of sight is made oblique ( $40^\circ$  w.r.t. the horizontal plane). Three respiratory signals was simultaneously tracked. Despite the small amplitudes of the thoracic elevations (5 mm), the protocol was able to detect some sagittal asymmetries. However, issuing an 1DoF elevation measurement per spot is still a modest objective w.r.t. to a protocol providing reliable 6DoF real-time measurements per spot. Moreover, if needed, the root shape tracking approach enables to perform more localized meas-



**Fig. 6.** Gating of respiratory motions through three thorax virtual spots. View (a) displays instance  $t_0$  and the static anatomical FoR of which two axes are parallel to the rest plane. View (b) superposes the consecutive locations of the FoRs linked to the three virtual spots. For didactic purpose only, this figure also displays the consensus map of the spotted shapes. View (c) depicts the corresponding elevation curves. The relative instantaneous elevations are defined through the projections of the corresponding FoR translation vectors on the vertical axis of the anatomical FoR. In this respiration mode, the highest thorax amplitude is 5 mm. This is to compare with the standard deviation of the filtered depth measurements still amounting to 5 mm.

urements. Therefore, upon adding a few additional virtual spotlights, this protocol would enable, through spatial interpolation, to provide a reliable 4D continuous elastic field describing thorax deformations.

## V. CONCLUSION AND PERSPECTIVES

To foster the significance of the RS concept, a.w.a. the effectiveness of the RST algorithm, two RS-based innovative protocols are proposed. The first protocol focuses on RS building tasks whereas the second focusses on RS tracking tasks. They corresponds to the two faces of the same algorithmic problem.

The first protocol, drawn from availability of cortical bone samples, shows how their common RS, together with its consensus map, can be made to operate as the core element of a transversal prosthesis-oriented design procedure that usually involves multidisciplinary design constraints.

The second protocol addresses in vivo kinematics measurements and operates through markerless tracking of the RS of consecutive skin shape instances. Moreover, while primarily focussing on in vivo scapula movements, this protocol puts a particular emphasis on accurate measurements of movements through a low-cost miniature 3D-video-camera. Although it involves a very compact equipment operating in conventional offices, its promising results lead to expect challenging measurements issued from high-end and expensive professional equipments in near future. However, due to the lack of precision of the 3D-video-cameras currently available, the *fully* markerless shape-based protocol could not be successfully implemented and, consequently, performing STA-free measurements of the scapula movements remains an open

problem. Thus, the full assessment of this new kinematics-oriented protocol remains to be done on other articular segments whereby well-proven alternative methods become operational – this work is in progress. However, as its successfully addresses respiratory gating through skin motions, this protocol already proves to be versatile.

The current convergence criterion of the RST algorithm may still appear difficult to meet and may thus lead to computational overhead. More generally, to deal efficiently with a very large number of instances, a technical improvement is still needed to lessen the computational demand of the core of the RST algorithm. Meanwhile, unlike gesture recognition tasks, medical-oriented objectives, for the most part, do not require to perform real-time analysis of the real-time acquisitions. Therefore, w.r.t. both accuracy and robustness improvements, the computational load of the RST algorithm should not be an impediment. Furthermore, powerful vectorial GPUs a.w.a. multi-core CPUs are now standard hardware components of laptop computers. Thus, bridging between these computational resources and algorithms will not remain an expert-wise task. Indeed, as of this writing, both GPU programming and multithreading software resources are on the way of standardization [20]. Therefore, as the RST algorithm is *intrinsically parallel*, it could become much more reactive without having to perform ad hoc approximations of the algorithm.

## REFERENCES

- [1] J.J. Jacq, T. Cresson, V. Burdin, C. Roux, “Performing accurate joint kinematics from 3D in vivo image sequences through consensus-driven simultaneous registration”, *IEEE Trans. on Biomedical Engineering*, Vol. 55, No. 5, May 2008, pp. 1620-1633
- [2] H. Hoppe, T. DeRose, T. Duchamp, J. McDonald, and W. Stuetzle. “Surface reconstruction from unorganized points”, *SIGGRAPH, Computer Graphics*, Vol. 26, pp. 71–78, 1992.
- [3] B. Curless and M. Levoy, “A volumetric method for building complex models from range images”, *SIGGRAPH*, pp. 303–312, 1996.
- [4] F.L. Bookstein, “Principal warps: thin-plate splines and the decomposition of deformations”, *IEEE trans. on Pattern Analysis and Machine Intelligence*, Vol. 11, No 6, 1989, pp. 567-585
- [5] F.L. Bookstein, “*Morphometric tools for landmark data, Geometry and Biology*”, Cambridge University Press, 1991
- [6] S. Kolta, A. Le Bras, D. Mitton, V. Bousson, J. A. de Guise, J. Fechtenbaum, J. D. Laredo, C. Roux, and W. Skalli, “Three-dimensional X-ray absorptiometry (3D-XA): a method for reconstruction of human bones using a dual X-ray absorptiometry device”, *Osteoporosis International*, Vol. 16, No 8, Aug. 2005, pp. 969-976
- [7] A. Leardini, L. Chiari, U. Della Croce, A. Cappozzo, “Human movement analysis using stereophotogrammetry Part 3. Soft tissue artefact assessment and compensation”, *Gait and Posture*, Vol. 21, 2005, pp. 212-225
- [8] K. Matsui, K. Shimada, P.D Andrew, “Deviation of skin marker from bone target during movement of the scapula”, *Journal of Orthopaedic Science*, Vol. 11, No. 2, Mar. 2006, pp. 180-184
- [9] S. Corazza, L. Mündermann, T. Andriacchi, “A framework for the functional identification of joint centers using markerless motion capture, validation for the hip joint”, *Journal of Biomechanics*, No. 40, Aug. 2007, pp. 3510–3515
- [10] G. Wu, F.C.T. van der Helm, H.E.J. Veeger, M. Makhsous, P. Van Roy, C. Anglin, J. Nagels, A.R. Karduna, K. McQuade, X. Wang, F.W. Werner, B. Buchholz, “ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion—Part II: shoulder, elbow, wrist and hand”, *Journal of Biomechanics*, Vol. 38, 2005, pp. 981–992
- [11] C. Schwartz, M. Lempereur, V. Burdin, J.J. Jacq, O. Rémy-Néris, “Shoulder motion analysis using simultaneous skin shape registration”, *Proc. of EMBC'07*, pp. 533-536, 23-26 Aug. 2007, Lyon, France
- [12] T. Weise, B. Leibe, L. Van Gool, “Fast 3D scanning with automatic motion compensation”, *CVPR*, Jul. 2007, pp. 1-8.

- [13] G. Yahav, G.J. Iddan, D. Mandelboum, “3D imaging camera for gaming application”, *In Proc. of Int. Conference on Consumer Electronics, Digest of Technical Papers*, 10-14 Jan. 2007, LasVegas, pp. 1-2
- [14] T. Oggier, F. Lustenberger, N. Blanc, “Miniature 3D TOF camera for real-time imaging”, *LNCS, Perception and Interactive Technologies*, LNAI 4021, Springer, 2006, pp. 212–216
- [15] S.A. Gudmundsson, H. Aanaes, R. Larsen, “Environmental effects on measurement uncertainties of Time-of-Flight cameras”, *In proc. of IEEE Int. Symposium on Signals, Circuits and Systems*, Romania, 13-14 Jul. 2007, pp. 1-4
- [16] J.J. Jacq, “Building and tracking root shapes – Application to biomechanics (I) : Estimating shoulder movements using a 3D-camera”, Institut TELECOM, TELECOM Bretagne, LaTIM, INSERM U650, 14 Jul. 2008, Mpeg-4 video, 1’50, 11 Mo (file currently annexed to this manuscript)
- [17] M. Böhme, M. Haker, T. Martinetz, E. Barth, “Shading constraint improves accuracy of Time-of-Flight measurements”, *Proc. Workshop on Time of Flight Camera based Computer Vision (TOF-CV), CVPR*, Jun. 28, 2008, Anchorage, pp. 1–6
- [18] B. Büttgen, H.A. El Mechat, F. Lustenberger, P. Seitz, “Pseudonoise optical modulation for real-time 3-D imaging with minimum interference”, *IEEE Trans. on Circuits and Systems*, Vol. 54, No. 10, Oct. 2007, pp. 2109-2119
- [19] C. Schaller, J. Penne, J. Hornegger, “Time-of-flight sensor for respiratory motion gating”, *Med. Phys.* 35 (7), Jul. 2008, pp. 3090-3093
- [20] A. Munshi, “OpenCL: parallel computing on the GPU and CPU”, *SIG-GRAPH’08 Course - Beyond Programmable Shading*, Aug. 14 2008, Los Angeles



**Romain Gérard**



**Christian Lefèvre**



**Olivier Rémy-Néris**



**Jean-José Jacq**



**Cédric Schwartz** received the Ph.D. degree in \_\_\_\_\_ at the University of Bretagne Occidentale within the Laboratoire de Traitement de l’Information Médicale, LaTIM-Inserm U650, France, in 2009. He is currently with \_\_\_\_\_ . His research interests are \_\_\_\_\_



**Christian Roux** (M’84–SM’95–F’05) received the Agregation degree in physics from the Ecole Normale Supérieure, Cachan, France, in 1978, and the Ph.D. degree from the Institut National Polytechnique, Grenoble, France, in 1980.

He joined the Institut TELECOM, TELECOM Bretagne, Brest, France, in 1982, became an Associate Professor in 1987, and then a Professor from 1987 onwards. He has been a Visiting Professor with the Medical Image Processing Group, Department of Radiology, University of Pennsylvania during 1992–1993, and a Distinguished International Research Fellow with the Department of Electrical Engineering, University of Calgary, Calgary, AB, Canada, in 1996 and 2003. He is currently the Head of the Laboratoire de Traitement de l’Information Médicale (LaTIM), INSERM U650, Brest. He is the author or coauthor of more than 100 papers, the author of four book chapters, and has edited three books. He is the holder of two patents. His current research interests include advanced medical information processing, and spatial and functional information modeling and analysis in medical images, with applications in various medical domains including orthopedics, gastroenterology, ophthalmology, and cardiology.

Prof. Roux was an Associate Editor for the IEEE TRANSACTIONS ON MEDICAL IMAGING during 1993–2000, and is a member of the Editorial board of the IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY and of the PROCEEDINGS OF THE IEEE.



**Valérie Burdin** (M’98) received the M.S. degree in applied mathematics from the University of Saint-Etienne, Saint-Etienne, France, in 1986, and the Ph.D. degree in telecommunications and signal processing from the University of Rennes I, Rennes, France, in 1992.

Since 1997, she has been a Senior Lecturer in the Department of Image and Information Processing, Institut TELECOM, TELECOM Bretagne, Brest, France, where she is engaged in research projects in medical image processing at the INSERM U650 Laboratory (LaTIM). Her current research interests include motion analysis and 3-D modeling of musculoskeletal systems.

Dr. Burdin is a member of the IEEE Engineering in Medicine and Biology (EMB) Society and the French Society of Biomedical Engineering (SFGMB). She is engaged in the organization of the IEEE EMBS International Summer School on Biomedical Imaging, held every 2 years at Berder Island, Brittany, France.

---

<sup>1</sup> The thorax ISB-FoR can then be defined from the vertical axis of the camera and the plane of the shoulders. The Scapula ISB-FoR and Humerus ISB-FoR are aligned with the Thorax ISB-FoR at time  $t_0$ . For didactic purpose only, the origin of these two FoRs are then approximatively shifted to 3D locations defined by their respective ISB's recommendations.

<sup>2</sup> Model SR-3100, MESA Imaging; phys. size: 50x67x42.3 mm; monocular sensor: QCIF 176x144 pix. This ToF model requires one 3D calibration file per modulation frequency ( $\nu = 19, 20, 21, 30$  MHz); depth encoding: 16 bits. Among models commercially available at the beginning of this work, this was the more suitable 3D-video-camera.

<sup>3</sup> The normalized quality of a triangle is defined as  $3^{3/4} 2 s^{1/2} l^{-1}$  where  $s$  (resp.  $l$ ) denotes its area (resp. perimeter). This coefficient expresses the ratio of its surface to its perimeter relative to this same ratio for an equilateral triangle with the same area.

<sup>4</sup> While conforming to ISB's recommendations, the curves depicted in Fig. 4.e show how difficult it is to get an intuitive account for general 3D rotations through a predetermined Euler's angles factorization for any 3DoF ball-and-socket joint. Indeed, an YZY Euler's factorization (i.e., subtracting  $90^\circ$  to YXY-Euler1 and adding  $90^\circ$  to YXY-Euler3 angles) should perform a more intuitive account of forward flexions of the arm.