IN PARTNERSHIP WITH:
CNRS

Institut polytechnique de Grenoble

Université Joseph Fourier (Grenoble)

Activity Report 2015

Project-Team MORPHEO

Capture and Analysis of Shapes in Motion

IN COLLABORATION WITH: Laboratoire Jean Kuntzmann (LJK)

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Vision, perception and multimedia interpretation

Table of contents

Project-Team MORPHEO

Creation of the Team: 2011 March 01, updated into Project-Team: 2014 January 01

Keywords:

Computer Science and Digital Science:

5.4.4. - 3D and spatio-temporal reconstruction
5.4.5. - Object tracking and motion analysis
5.5.4. - Animation
5.6. - Virtual reality, augmented reality

Other Research Topics and Application Domains:

9.2.2. - Cinema, Television
9.2.3. - Video games

1. Members

Research Scientists

Edmond Boyer [Team leader, Inria, Senior Researcher, HdR]
Julien Pansiot [Inria, Starting Research position]
Lionel Reveret [Inria, Researcher, HdR]
Stefanie Wuhrer [Inria, Researcher, from Feb 2015]

Faculty Members

Jean Sébastien Franco [Grenoble INP , Associate Professor]
Franck Hétroy-Wheeler [Grenoble INP, Associate Professor, HdR]

Engineers

Mickaël Heudre [Inria]
Thomas Pasquier [Inria, until Jun 2015]

PhD Students

Benjamin Allain [Inria]
Adnane Boukhayma [Inria, granted by Conseil Régional Rhône-Alpes]
Vincent Leroy [Inria, from Feb 2015]
Romain Rombourg [Univ. Grenoble Alpes, from Oct 2015]
Aurela Shehu [Saarland University, from Feb 2015]
Vagia Tsiminaki [Inria, granted by FP7 RE@CT project]
Li Wang [Univ. Grenoble I]
Jinlong Yang [Inria, from Oct 2015]
Benjamin Aupetit [Univ. Grenoble Alpes, until Jan 2015]

Post-Doctoral Fellow

Mohammad Rouhani [Inria, until March 2015]

Administrative Assistant

Elodie Toihein [Inria]

2. Overall Objectives

2.1. Overall Objectives

Morpheo's main objective is the ability to perceive and to interpret moving shapes using systems of multiple cameras for the analysis of animal motion, animation synthesis and immersive and interactive environments. Multiple camera systems allow dense information on both shapes and their motion to be recovered from visual cues. Such ability to perceive shapes in motion brings a rich domain for research investigations on how to model, understand and animate real dynamic shapes. In order to reach this objective, several scientific and technological challenges must be faced:

A first challenge is to be able to recover shape information from videos. Multiple camera setups allow to acquire shapes as well as their appearances with a reasonable level of precision. However most effective current approaches estimate static 3D shapes and the recovery of temporal information, such as motion, remains a challenging task. Another challenge in the acquisition process is the ability to handle heterogeneous sensors with different modalities as available nowadays: color cameras, time of flight cameras, stereo cameras and structured light scanners, etc.

A second challenge is the analysis of shapes. Few tools have been proposed for that purpose and recovering the intrinsic nature of shapes is an actual and active research domain. Of particular interest is the study of animal shapes and of their associated articulated structures. An important task is to automatically infer such properties from temporal sequences of 3D models as obtained with the previously mentioned acquisition systems. Another task is to build models for classes of shapes, such as animal species, that allow for both shape and pose variations.

A third challenge concerns the analysis of the motion of shapes that move and evolve, typically humans. This has been an area of interest for decades and the challenging innovation is to consider for this purpose dense motion fields, obtained from temporally consistent 3D models, instead of traditional sparse point trajectories obtained by tracking particular features on shapes, e.g. motion capture systems. The interest is to provide full information on both motions and shapes and the ability to correlate these information.The main tasks that arise in this context are first to find relevant indices to describe the dynamic evolutions of shapes and second to build compact representations for classes of movements.

A fourth challenge tackled by Morpheo is immersive and interactive systems. Such systems rely on real time modeling, either for shapes, motion or actions. Most methods of shape and motion retrieval turn out to be fairly complex, and quickly topple hardware processing or bandwidth limitations, even with a limited number of cameras. Achieving interactivity thus calls for scalable methods and research of specific distribution and parallelization strategies.

3. Research Program

3.1. Shape Acquisition

Multiple camera setups allow to acquire shapes, i.e. geometry, as well as their appearances, i.e. photometry, with a reasonable level of precision. However fundamental limitations still exist, in particular today's state-of-the-art approaches do not fully exploit the redundancy of information over temporal sequences of visual observations. Despite an increasing interest of the computer vision communities in the past years, the problem is still far from solved other than in specific situations with restrictive assumptions and configurations. Our goal in this research axis is to open the acquisition process to more general assumptions, e.g. no specific lighting or background conditions, scenes with evolving topologies, , and fully leverage temporal aspects of the acquisition process.

3.2. Bayesian Inference

Acquisition of 4D Models can often be conveniently formulated as a Bayesian estimation or learning problem. Various generative and graphical models can be proposed for the problems of shape and appearance modeling over time sequences, and motion segmentation. The idea of these generative models is to predict the noisy measurements (e.g. pixel values, measured 3D points or speed quantities) from a set of parameters describing the unobserved scene state (e.g. shape and appearance), which in turn can be estimated using Bayes' rule to solve the inverse problem. The advantages of this type of modeling are numerous, as they enable to model the noisy relationships between observed and unknown quantities specific to the problem, deal with outliers, and allow to efficiently account for various types of priors about the scene and its semantics. Sensor models for different modalities can also easily be seamlessly integrated and jointly used, which remains central to our goals.

Since the acquisition problems often involve a large number of variables, a key challenge is to exhibit models which correctly account for the observed phenoma, while keeping reasonable estimation times, sometimes with a real-time objective. Maximum likelihood / maximum a posteriori estimation and approximate inference techniques, such as Expectation Maximization, Variational Bayesian inference, or Belief Propagation, are useful tools to keep the estimation tractable. While 3D acquisition has been extensively explored, the research community faces many open challenges in how to model and specify more efficient priors for 4D acquisition and temporal evolution.

3.3. Shape Analysis

Shape analysis has received much attention from the scientific community and recovering the intrinsic nature of shapes is currently an active research domain. Of particular interest is the study of human and animal shapes and their associated articulated underlying structures, i.e. skeletons, since applications are numerous, either in the entertainment industry or for medical applications, among others. Our main goals in this research axis are : the understanding of a shape's global structure, and a pose-independent classification of shapes.

3.4. Shape Tracking

Recovering the temporal evolution of a deformable surface is a fundamental task in computer vision, with a large variety of applications ranging from the motion capture of articulated shapes, such as human bodies, to the deformation of complex surfaces such as clothes. Methods that solve for this problem usually infer surface evolutions from motion or geometric cues. This information can be provided by motion capture systems or one of the numerous available static 3D acquisition modalities. In this inference, methods are faced with the challenging estimation of the time-consistent deformation of a surface from cues that can be sparse and noisy. Such an estimation is an ill posed problem that requires prior knowledge on the deformation to be introduced in order to limit the range of possible solutions. Our goal is to devise robust and accurate solutions based on new deformation models that fully exploit the geometric and photometric information available.

3.5. Dynamic Motion Modeling

Multiple views systems can significantly change the paradigm of motion capture. Traditional motion capture systems provide 3D trajectories of a sparse set of markers fixed on the subject. These trajectories can be transformed into motion parameters on articulated limbs with the help of prior models of the skeletal structure. However, such skeletal models are mainly robotical abstractions that do not describe the true morphology and anatomical motions of humans and animals. On the other hand, 4D models (temporally consistent mesh sequences) provide dense motion information on body's shape while requiring less prior assumption. They represent therefore a new rich source of information on human and animal shape movements. The analysis of such data has already received some attention but most existing works model motion through static poses and do not consider yet dynamic information. Such information (e.g. trajectories and speed) is anyway required to analyse walking or running sequences. We will investigate this research direction with the aim to propose and study new dynamic models.

4. Application Domains

4.1. 4D modeling

Modeling shapes that evolve over time, analyzing and interpreting their motion has been a subject of increasing interest of many research communities including the computer vision, the computer graphics and the medical imaging communities. Recent evolutions in acquisition technologies including 3D depth cameras (Time-of-Flight and Kinect), multi-camera systems, marker based motion capture systems, ultrasound and CT scans have made those communities consider capturing the real scene and their dynamics, create 4D spatio-temporal models, analyze and interpret them. A number of applications including dense motion capture, dynamic shape modeling and animation, temporally consistent 3D reconstruction, motion analyzes and interpretation have therefore emerged.

4.2. Shape Analysis

Most existing shape analysis tools are local, in the sense that they give local insight about an object's geometry or purpose. The use of both geometry and motion cues makes it possible to recover more global information, in order to get extensive knowledge about a shape. For instance, motion can help to decompose a 3D model of a character into semantically significant parts, such as legs, arms, torso and head. Possible applications of such high-level shape understanding include accurate feature computation, comparison between models to detect defects or medical pathologies, and the design of new biometric models or new anthropometric datasets.

4.3. Human Motion Analysis

The recovery of dense motion information enables the combined analyses of shapes and their motions. Typical examples include the estimation of mean shapes given a set of 3D models or the identification of abnormal deformations of a shape given its typical evolutions. The interest arises in several application domains where temporal surface deformations need to be captured and analysed. It includes human body analyses for which potential applications are anyway numerous and important, from the identification of pathologies to the design of new prostheses.

4.4. Interaction

The ability to build models of humans in real time allows to develop interactive applications where users interact with virtual worlds. The recent Kinect proposed by Microsoft illustrates this principle with game applications using human inputs perceived with a depth camera. Other examples include gesture interfaces using visual inputs. A challenging issue in this domain is the ability to capture complex scenes in natural environments. Multi-modal visual perception, e.g. depth and color cameras, is one objective in that respect.

5. Highlights of the Year

5.1. Highlights of the Year

- The multi-camera platform Kinovis (http://kinovis.inrialpes.fr) was inaugurated on May 26th 2015. Kinovis is French *Equipement d'excellence* (Equipex project) that provides a unique acquisition platform with 68 color cameras and enables therefore high precision 4D modeling of dynamic scenes.
- The QuickCSG boolean mesh computation software developed within the context of the Kinovis platform was transferred in November of 2015, to a (contractually undisclosed) major industrial actor of the 3D business.

6. New Software and Platforms

6.1. 4D repository

FUNCTIONAL DESCRIPTION

This website hosts dynamic mesh sequences reconstructed from images captured using a multi-camera set up. Such mesh-sequences offer a new promising vision of virtual reality, by capturing real actors and their interactions. The texture information is trivially mapped to the reconstructed geometry, by back-projecting from the images. These sequences can be seen from arbitrary viewing angles as the user navigates in 4D (3D geometry + time) . Different sequences of human / non-human interaction can be browsed and downloaded from the data section.

- Contact: Bruno Raffin
- URL: http://4drepository.inrialpes.fr/

6.2. ETHOMICE

KEYWORDS: Biology - Health - Biomechanics - Motion analysis - Ethology - Mouse
FUNCTIONAL DESCRIPTION

Ethomice is a motion analysis software to characterize motor behavior of small vertebrates such as mice or rats. From a multiple views video input, a biomechanical model of the skeleton is registered. Study on animal model is the first important step in Biology and Clinical research. In this context, the analysis of the neuro-motor behaviour is a frequent cue to test the effect of a gene or a drug. Ethomice is a platform for simulation and analysis of the small laboratory animal, such as rat or mouse. This platform links the internal skeletal structure with 3D measurements of the external appearance of the animal under study. From a stream of multiple views video, the platform aims at delivering a three dimensional analysis of the body posture and the behaviour of the animal.

- Participants: Lionel Reveret
- Partners: CNRS - Inria - Université Descartes - ICS
- Contact: Lionel Reveret
- URL: http://morpheo.inrialpes.fr/people/reveret/ethomice

6.3. Lucy Viewer

FUNCTIONAL DESCRIPTION

Lucy Viewer is an interactive viewing software for 4D models, i.e, dynamic three-dimensional scenes that evolve over time. Each 4D model is a sequence of meshes with associated texture information, in terms of images captured from multiple cameras at each frame.

- Participants: Edmond Boyer and Florent Lagaye
- Contact: Edmond Boyer
- URL: http://4drepository.inrialpes.fr/lucy_viewer/

6.4. QuickCSG

KEYWORDS: 3D modeling - CAD - 3D reconstruction - Geometric algorithms
FUNCTIONAL DESCRIPTION

QuickCSG is a library and command-line application that computes boolean operations between polyhedra. It is able to direcly compute resulting solids from an arbitrary number of inputs and for an arbitrary boolean combination function, with state of the art execution times.

- Participants: Matthys Douze, Jean-Sébastien Franco and Bruno Raffin
- Partner: INP Grenoble
- Contact: Matthys Douze
- URL: http://kinovis.inrialpes.fr/static/QuickCSG/

6.5. Shape Tracking

FUNCTIONAL DESCRIPTION

We are developing a software suite to track shapes over temporal sequences. The motivation is to provide temporally coherent 4D Models, i.e. 3D models and their evolutions over time, as required by motion related applications such as motion analysis. This software takes as input a temporal sequence of 3D models in addition to a template and estimates the template deformations over the sequence that fit the observed 3D models. This software is particularly developed in the context of the FUI project Creamove.

- Contact: Edmond Boyer

6.6. Platforms

6.6.1. Platform Kinovis

Kinovis (http://kinovis.inrialpes.fr/) is a multi-camera acquisition project that was was selected within the call for proposals ”Equipements d’Excellence” of the program “Investissement d’Avenir˝ funded by the French government. The project involves 2 institutes: the Inria Grenoble Rhône-Alpes, the université Joseph Fourier and 4 laboratories: the LJK (laboratoire Jean Kuntzmann - applied mathematics), the LIG (laboratoire d’informatique de Grenoble - Computer Science), the Gipsa lab (Signal, Speech and Image processing) and the LADAF (Grenoble Hospitals - Anatomy). The Kinovis environment is composed of 2 complementary platforms. A first platform located at Inria Grenoble with a 10mx10m acquisition surface is equipped with 68 color cameras and 20 IR motion capture (mocap) cameras. It is the evolution of the Grimage platform towards the production of better models of more complex dynamic scenes. A second platform located at Grenoble Hospitals, within the LADAF anatomy laboratory, is equipped with 10 color and 2 X-ray cameras to enable combined analysis of internal and external shape structures, typically skeleton and bodies of animals. Installation works of both platforms started in 2013 and are now finished. Both platforms have already demonstrated their potential through a range of projects lead by the team and externally. Members of Morpheo are highly involved in this project. Edmond Boyer is coordinating this project and Lionel Reveret is in charge of the LADAF platform. Thomas Pasquier, Mickaël Heudre and Julien Pansiot are managing the technical resources of both platforms.

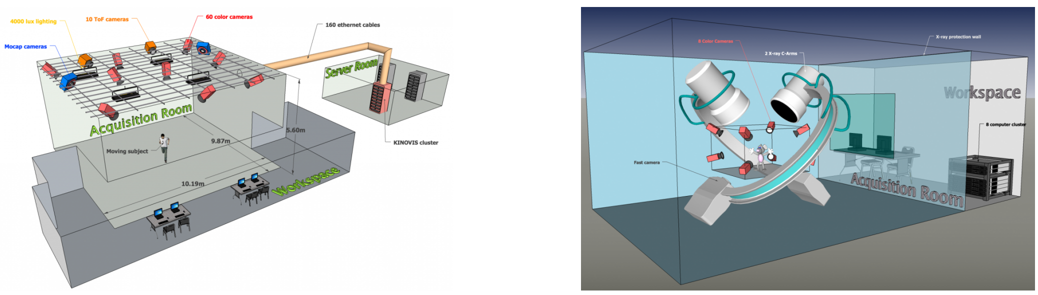

Figure 1. Kinovis platforms: on the left the Inria platform; on the right Grenoble Hospital platform.

6.6.2. Multicamera platform for video analysis of mice behavior

This project is a follow-up of the experimental set-up developed for a CNES project with Mathieu Beraneck from the CESeM laboratory (centre for the study of sensorimotor control, CNRS UMR 8194) at the Paris-Descartes University. The goal of this project was to analyze the 3D body postures of mice with various vestibular deficiencies in low gravity condition (3D posturography) during a parabolic flight campaign. The set-up has been now adapted for new experiments on motor-control disorders for other mice models.

This experimental platform is currently under development for a broader deployment for high throughput phenotyping with the technology transfer project ETHOMICE. This project involves a close relationship with the CESeM laboratory and the European Mouse Clinical Institute in Strasbourg (Institut Clinique de la Souris, ICS).

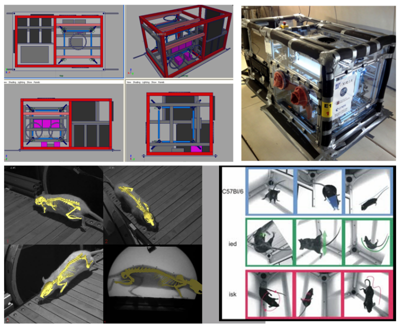

Figure 2. Ethomice: Experimental platform for video analysis of mice behavior.

7. New Results

7.1. QuickCSG: Arbitrary and Faster Boolean Combinations of N Solids

While studied over several decades, the computation of boolean operations on polyhedra is almost always addressed by focusing on the case of two polyhedra. For multiple input polyhedra and an arbitrary boolean operation to be applied, the operation is decomposed over a binary CSG tree, each node being processed separately in quasilinear time. For large trees, this is both error prone due to intermediate geometry and error accumulation, and inefficient because each node yields a specific overhead. We introduce a fundamentally new approach to polyhedral CSG evaluation, addressing the general N-polyhedron case. We propose a new vertex-centric view of the problem, which both simplifies the algorithm computing resulting geometric contributions, and vastly facilitates its spatial decomposition. We then embed the entire problem in a single KD-tree, specifically geared toward the final result by early pruning of any region of space not contributing to the final surface. This not only improves the robustness of the approach, it also gives it a fundamental speed advantage, with an output complexity depending on the output mesh size instead of the input size as with usual approaches. Complemented with a task-stealing parallelization, the algorithm achieves breakthrough performance, one to two orders of magnitude speedups with respect to state-of-the-art CPU algorithms, on boolean operations over two to several dozen polyhedra. The algorithm is also shown to outperform recent GPU implementations and approximate discretizations, while producing a topologically exact output without redundant facets. This algorithm was published as Inria research report [16].

7.2. An Efficient Volumetric Framework for Shape Tracking

Recovering 3D shape motion using visual information is an important problem with many applications in computer vision and computer graphics, among other domains. Most existing approaches rely on surface-based strategies, where surface models are fit to visual surface observations. While numerically plausible, this paradigm ignores the fact that the observed surfaces often delimit volumetric shapes, for which deformations are constrained by the volume inside the shape. Consequently, surface-based strategies can fail when the observations define several feasible surfaces, whereas volumetric considerations are more restrictive with

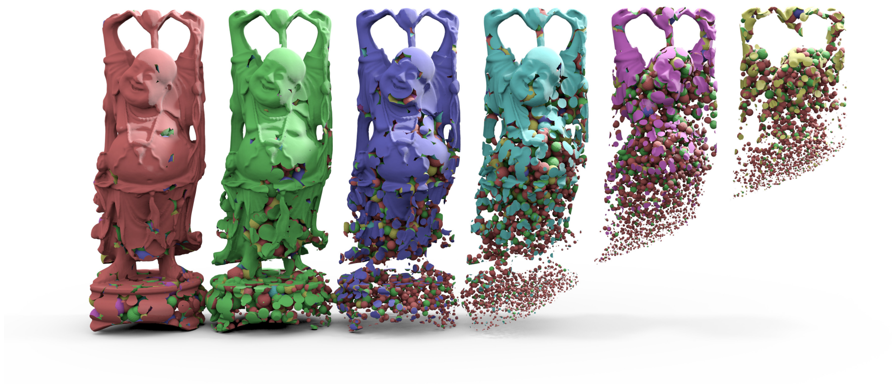

Figure 3. Intersection of 6 Buddhas with the union of 100,000 spheres (total 24 million triangles). Computed in 8 seconds on a desktop machine [16]

respect to the admissible solutions. In this work, we investigate a novel volumetric shape parametrization to track shapes over temporal sequences. In constrast to Eulerian grid discretizations of the observation space, such as voxels, we consider general shape tesselations yielding more convenient cell decompositions, in particular the Centroidal Voronoi Tesselation. With this shape representation, we devise a tracking method that exploits volumetric information, both for the data term evaluating observation conformity, and for expressing deformation constraints that enforce prior assumptions on motion. Experiments on several datasets demonstrate similar or improved precisions over state-of-the-art methods, as well as improved robustness, a critical issue when tracking sequentially over time frames. This work was accepted as **oral** at CVPR 2015 (less than 3% acceptance rate) [8].

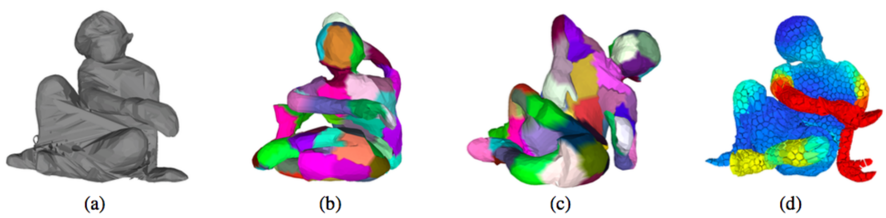

Figure 4. Frames of the GOALKEEPER *dataset acquired on the Kinovis platform. (a) Visual hull input. (b) Tracking result of Cagniart et al. 2010. (c) Allain et al. 2014. (d) This method [8]. Note the improved angular shapes and the improved robustness.*

7.3. Sparse Multi-View Consistency for Object Segmentation

Multiple view segmentation consists in segmenting objects simultaneously in several views. A key issue in that respect and compared to monocular settings is to ensure propagation of segmentation information between views while minimizing complexity and computational cost. In this work, we first investigate the idea that examining measurements at the projections of a sparse set of 3D points is sufficient to achieve this goal. The proposed algorithm softly assigns each of these 3D samples to the scene background if it projects on the background region in at least one view, or to the foreground if it projects on foreground region in all views. Second, we show how other modalities such as depth may be seamlessly integrated in the model and benefit the segmentation. The paper exposes a detailed set of experiments used to validate the algorithm, showing results comparable with the state of art, with reduced computational complexity. We also discuss the use of different modalities for specific situations, such as dealing with a low number of viewpoints or a scene with color ambiguities between foreground and background. This work was published as article in the PAMI journal [3].

Figure 5. Three views of the PLANT *dataset as processed by our method for mutli-view silhouette extraction [3].*

7.4. Building Statistical Shape Spaces for 3D Human Modeling

Statistical models of 3D human shape and pose learned from scan databases have developed into valuable tools to solve a variety of vision and graphics problems. Unfortunately, most publicly available models are of limited expressiveness as they were learned on very small databases that hardly reflect the true variety in human body shapes. In this paper, we contribute by rebuilding a widely used statistical body representation from the largest commercially available scan database, and making the resulting model available to the community (visit http://humanshape.mpi-inf.mpg.de). As preprocessing several thousand scans for learning the model is a challenge in itself, we contribute by developing robust best practice solutions for scan alignment that quantitatively lead to the best learned models. We make implementations of these preprocessing steps also publicly available. We extensively evaluate the improved accuracy and generality of our new model, and show its improved performance for human body reconstruction from sparse input data. This work was published as Max Planck research report [17].

7.5. A Groupwise Multilinear Correspondence Optimization for 3D Faces

Multilinear face models are widely used to model the space of human faces with expressions. For databases of 3D human faces of different identities performing multiple expressions, these statistical shape models decouple identity and expression variations. To compute a high-quality multilinear face model, the quality of the registration of the database of 3D face scans used for training is essential. Meanwhile, a multilinear face model can be used as an effective prior to register 3D face scans, which are typically noisy and incomplete. Inspired by the minimum description length approach, we propose the first method to jointly optimize a multilinear model and the registration of the 3D scans used for training. Given an initial registration, our approach fully automatically improves the registration by optimizing an objective function that measures the compactness of the multilinear model, resulting in a sparse model. We choose a continuous representation for each face shape that allows to use a quasi-Newton method in parameter space for optimization. We show that

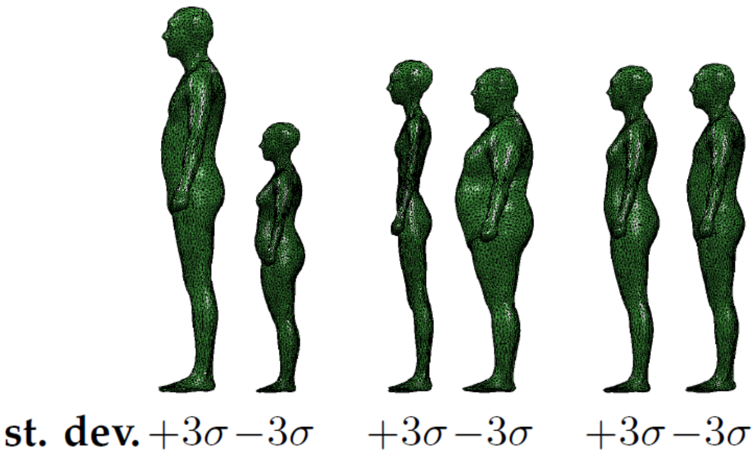

Figure 6. Visualization of the first three principal components learned from a large database of posture-normalized 3D human body scans [17].

our approach is computationally significantly more efficient and leads to correspondences of higher quality than existing methods based on linear statistical models. This allows us to evaluate our approach on large standard 3D face databases and in the presence of noisy initializations. This work was published at the ICCV conference [9].

7.6. A statistical shape space model of the palate surface trained on 3D MRI scans of the vocal tract

We describe a minimally-supervised method for computing a statistical shape space model of the palate surface. The model is created from a corpus of volumetric magnetic resonance imaging (MRI) scans collected from 12 speakers. We extract a 3D mesh of the palate from each speaker, then train the model using principal component analysis (PCA). The palate model is then tested using 3D MRI from another corpus and evaluated using a high-resolution optical scan. We find that the error is low even when only a handful of measured coordinates are available. In both cases, our approach yields promising results. It can be applied to extract the palate shape from MRI data, and could be useful to other analysis modalities, such as electromagnetic articulography (EMA) and ultrasound tongue imaging (UTI). This work was published at the 18th International Congress of Phonetic Sciences [11].

7.7. Toward User-specific Tracking by Detection of Human Shapes in Multi-Cameras

Human shape tracking consists in fitting a template model to temporal sequences of visual observations. It usually comprises an association step, that finds correspondences between the model and the input data, and a deformation step, that fits the model to the observations given correspondences. Most current approaches find their common ground with the Iterative-Closest-Point (ICP) algorithm, which facilitates the association step with local distance considerations. It fails when large deformations occur, and errors in the association tend to propagate over time. In this paper, we propose a discriminative alternative for the association, that leverages random forests to infer correspondences in one shot. It allows for large deformations and prevents tracking errors from accumulating. The approach is successfully integrated to a surface tracking framework that recovers human shapes and poses jointly. When combined with ICP, this discriminative association proves to yield better accuracy in registration, more stability when tracking over time, and faster convergence. Evaluations on existing datasets demonstrate the benefits with respect to the state-of-the-art. This work was published at CVPR 2015 [12].

7.8. Video based Animation Synthesis with the Essential Graph

We propose a method to generate animations using video-based mesh sequences of elementary movements of a shape. New motions that satisfy high-level user-specified constraints are built by recombining and interpolating the frames in the observed mesh sequences. The interest of video based meshes is to provide real full shape information and to enable therefore realistic shape animations. A resulting issue lies, however, in the difficulty to combine and interpolate human poses without a parametric pose model, as with skeleton based animations. To address this issue, our method brings two innovations that contribute at different levels: Locally between two motion sequences, we introduce a new approach to generate realistic transitions using dynamic time warping; More globally, over a set of motion sequences, we propose the essential graph as an efficient structure to encode the most realistic transitions between all pairs of input shape poses. Graph search in the essential graph allows then to generate realistic motions that are optimal with respect to various user-defined constraints. We present both quantitative and qualitative results on various 3D video datasets. They show that our approach compares favourably with previous strategies in this field that use the motion graph. This work was published at the 3DV 2015 conference [10].

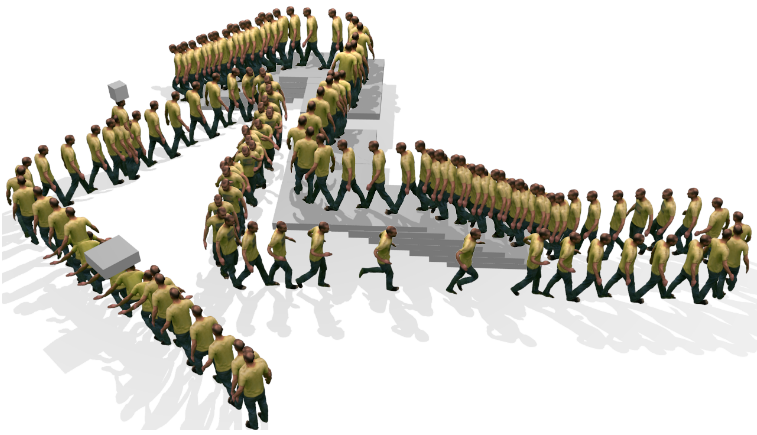

Figure 7. Example of 4D animation generated using by combining recorded 4D sequences [10].

7.9. Implicit B-Spline Surface Reconstruction

This paper presents a fast and flexible curve/surface reconstruction technique based on implicit b-spline. This representation does not require any parameterization and it is locally supported. This fact has been exploited in this paper to propose a reconstruction technique through solving a sparse system of equations. This method is further accelerated to reduce the dimension to the active control lattice. Moreover, the surface smoothness and user interaction are allowed for controlling the surface. Finally, a novel weighting technique has been introduced in order to blend small patches and smooth them in the overlapping regions. The whole framework is very fast and efficient and can handle large cloud of points with low computational cost. The experimental results show the flexibility and accuracy of the proposed algorithm to describe objects with complex topologies. Comparisons with other fitting methods highlight the superiority of the proposed approach in the presence of noise and missing data. This work was published as journal article in IEEE Transactions on Image Processing [6].

7.10. A Bayesian Approach to Multi-view 4D Modeling

This paper considers the problem of automatically recovering temporally consistent animated 3D models of arbitrary shapes in multi-camera setups. An approach is presented that takes as input a sequence of frame-wise reconstructed surfaces and iteratively deforms a reference surface such that it fits the input observations. This approach addresses several issues in this field that include: large frame-to-frame deformations, noise, missing data, outliers and shapes composed of multiple components with arbitrary geometries. The problem is cast as a geometric registration with two major features. First, surface deformations are modeled using mesh decomposition into elements called patches. This strategy ensures robustness by enabling flexible regularization priors through inter-patch rigidity constraints. Second, registration is formulated as a Bayesian estimation that alternates between probabilistic datal-model association and deformation parameter estimation. This accounts for uncertainties in the acquisition process and allows for noise, outliers and missing geometries in the observed meshes. In the case of marker-less 3D human motion capture, this framework can be specialized further with additional articulated motion constraints. Extensive experiments on various 4D datasets show that complex scenes with multiple objects of arbitrary nature can be processed in a robust way. They also demonstrate that the framework can capture human motion and provides visually convincing as well as quantitatively reliable human poses. This work was published as journal article in International Journal on Computer Vision (IJCV) [4].

7.11. A Hierarchical Approach for Regular Centroidal Voronoi Tessellations

In this paper we consider Centroidal Voronoi Tessellations (CVTs) and study their regularity. CVTs are geometric structures that enable regular tessellations of geometric objects and are widely used in shape modeling and analysis. While several efficient iterative schemes, with defined local convergence properties, have been proposed to compute CVTs, little attention has been paid to the evaluation of the resulting cell decompositions. In this paper, we propose a regularity criterion that allows us to evaluate and compare CVTs independently of their sizes and of their cell numbers. This criterion allows us to compare CVTs on a common basis. It builds on earlier theoretical work showing that second moments of cells converge to a lower bound when optimising CVTs. In addition to proposing a regularity criterion, this paper also considers computational strategies to determine regular CVTs. We introduce a hierarchical framework that propagates regularity over decomposition levels and hence provides CVTs with provably better regularities than existing methods. We illustrate these principles with a wide range of experiments on synthetic and real models.

This work was published as a journal article in Computer Graphics Forum [7].

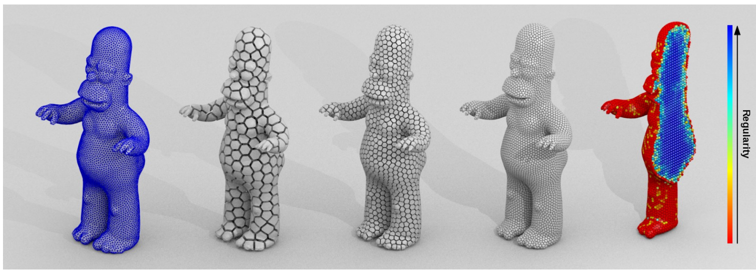

Figure 8. Hierarchical computation of a centroidal Voronoi tessellation from a 3D mesh [7]. Inside cells are very regular.

7.12. Just Noticeable Distortion Profile for Flat-Shaded 3D Mesh Surfaces

It is common that a 3D mesh undergoes some lossy operations (e.g., compression, watermarking and transmission through noisy channels), which can introduce geometric distortions as a change in vertex position. In most cases the end users of 3D meshes are human beings; therefore, it is important to evaluate the visibility of introduced vertex displacement. In this paper we present a model for computing a Just Noticeable Distortion (JND) profile for flat-shaded 3D meshes. The proposed model is based on an experimental study of the properties of the human visual system while observing a flat-shaded 3D mesh surface, in particular the contrast sensitivity function and contrast masking. We first define appropriate local perceptual properties on 3D meshes. We then detail the results of a series of psychophysical experiments where we have measured the threshold needed for a human observer to detect the change in vertex position. These results allow us to compute the JND profile for flat-shaded 3D meshes. The proposed JND model has been evaluated via a subjective experiment, and applied to guide 3D mesh simplification as well as to determine the optimal vertex coordinates quantization level for a 3D model.

This work was published as a journal article in IEEE Transactions on Visualization and Computer Graphics [5].

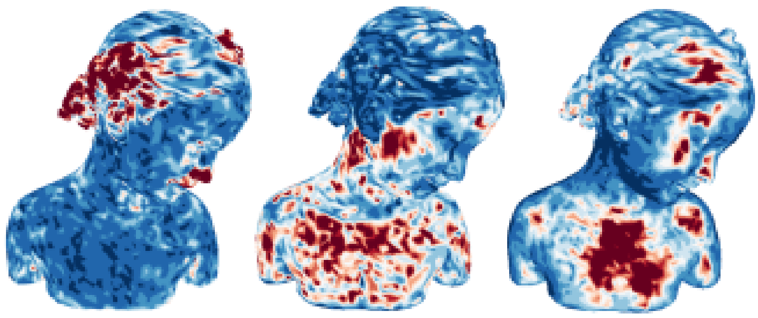

Figure 9. Just noticeable distortion profile in a light independent mode (left, middle) or with a light fixed in front of the model (right), for vertex displacements in the normal direction (left, right) or in the tangent direction (middle) [5].

8. Bilateral Contracts and Grants with Industry

8.1. QuickCSG Contract with undisclosed industrial partner

QuickCSG software was licensed in october 2015 to an industrial partner whose name is contractually kept undisclosed for a finite time period. QuickCSG is being integrated into the partner's software and is scheduled to be sold with this industrial partner's products during the year of 2016. An additional support contract has been signed with this partner for the purpose of the transfer.

9. Partnerships and Cooperations

9.1. ARC6 project PADME – Perceptual quality Assessment of Dynamic MEshes and its applications

In this project, we propose to use a new and experimental "bottom-up" approach to study an interdisciplinary problem, namely the objective perceptual quality assessment of 3D dynamic meshes (i.e., shapes in motion with temporal coherence). The objectives of the proposed project are threefold:

1. to understand the HVS (human visual system) features when observing 3D animated meshes, through a series of psychophysical experiments;
2. to develop an efficient and open-source objective quality metric for dynamic meshes based on the results of the above experiments;
3. to apply the learned HVS features and the derived metric to the application of compression and/or watermarking of animated meshes.

This work is funded by the Rhône-Alpes région through an ARC6 grant for the period 2013-2016. The three partners are LIRIS (University Lyon 1, Florent Dupont), GIPSA-Lab (CNRS, Kai Wang) and LJK (University of Grenoble, Franck Hétroy-Wheeler). A PhD student, Georges Nader, is working on this project.

9.2. National Initiatives

9.2.1. Motion analysis of laboratory rodents

In order to evaluate the scalabililty of previous work on motion analysis of laboratory rodents, a collaboration has been initiated with the Institut Clinique de la Souris (ICS), in Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC). This institute is dedicated to phenotypying of mice and requires reliable motion analysis tools. A multicamera plateform has been deployed at ICS and will be exploited next year for tests ranging from one to two hundreds mice.

9.2.2. ANR

9.2.2.1. ANR project Achmov – Accurate Human Modeling in Videos

The technological advancements made over the past decade now allow the acquisition of vast amounts of visual information through the use of image capturing devices like digital cameras or camcorders. A central subject of interest in video are the humans, their motions, actions or expressions, the way they collaborate and communicate. Analyzing video data of humans, collected for complex real-world events–extracting high-fidelity content, transferring raw data into knowledge–, detecting, reconstructing or understanding human motion are problems of key importance for the advancement of a variety of technological fields, including video coding, entertainment, culture, animation and virtual reality, intelligent human-computer interfaces, protection and security. The visual analysis of humans in real-world environments, indoors and outdoors, faces major scientific and computational challenges however. The proportions of the human body vary largely across individuals, any single human body has many degrees of freedom due to articulations, and individual limbs deform due to moving muscles and clothing. Finally, real-world events involve multiple interacting humans occluded by each other or by other objects, and the scene conditions may also vary due to camera motion or lighting changes. All these factors make appropriate models of human structure, motion and action difficult to construct and difficult to estimate from images. The goal of ACHMOV is to extract detailed representations of multiple interacting humans in real-world environments in an integrated fashion through a synergy between detection, figure-ground segmentation and body part labeling, accurate 3D geometric methods for kinematic and shape modeling, and large-scale statistical learning techniques. By integrating the complementary expertise of two teams (one French, MORPHEO and one Romanian, CLVP), with solid prior track records in the field, there are considerable opportunities to move towards processing complex real world scenes of multiple interacting people, and be able to extract rich semantic representations with high fidelity. This would enable interpretation, recognition and synthesis at unprecedented levels of accuracy and in considerably more realistic setups than currently considered. This project was kicked off on November 26th, 2015, in Bucharest, Romania.

9.2.3. Competitivity Clusters

9.2.3.1. FUI project Creamove

Creamove is a collaboration between the Morpheo team of the Inria Grenoble Rhône-Alpes, the 4D View Solution company specialised in multi-camera acquisition systems, the SIP company specialised in multi-media and interactive applications and a choreographer. The objective is to develop new interactive and artistic applications where humans can interact in 3D with virtual characters built from real videos. Dancer performances will be pre-recorded in 3D and used on-line to design new movement sequences based on inputs coming from human bodies captured in real time. Website: http://www.creamove.fr.

9.3. European Initiatives

9.3.1. FP7 & H2020 Projects

9.3.1.1. Re@ct

Type: FP7 COOPERATION

Defi: IMMERSIVE PRODUCTION AND DELIVERY OF INTERACTIVE 3D CONTENT

Instrument: Specific Targeted Research Project

Objectif: Networked Media ans Search Systems

Duration: December 2011 - November 2014 (Evaluation January through March 2015)

Coordinator: BBC (UK)

Partner: BBC (UK), Fraunhofer HHI (Germany), University of Surrey (UK), Artefacto (France), OMG (UK).

Inria contact: Jean-Sébastien Franco, Edmond Boyer

Abstract:RE@CT will introduce a new production methodology to create film-quality interactive characters from 3D video capture of actor performance. Recent advances in graphics hardware have produced interactive video games with photo-realistic scenes. However, interactive characters still lack the visual appeal and subtle details of real actor performance as captured on film. In addition, existing production pipelines for authoring animated characters are highly labour intensive. RE@CT aims to revolutionise the production of realistic characters and significantly reduce costs by developing an automated process to extract and represent animated characters from actor performance capture in a multiple camera studio. The key innovation is the development of methods for analysis and representation of 3D video to allow reuse for real-time interactive animation. This will enable efficient authoring of interactive characters with video quality appearance and motion. The project builds on the latest advances in 3D and free-viewpoint video from the contributing project partners. For interactive applications, the technical challenges are to achieve another step change in visual quality and to transform captured 3D video data into a representation that can be used to synthesise new actions and is compatible with current gaming technology.

9.4. International Initiatives

9.4.1. Inria International Partners

9.4.1.1. Declared Inria International Partners

9.4.1.1.1. Joint projects with the Forestry Commission, UK

Two common works with an ecophysiologist from the British Forestry Commission, Eric Casella, are currently carried out. The first one aims at detecting, analysing and correcting acquisition noise from terrestrial laser scans (t-LiDAR) of plants and trees. The second one aims at reconstructing accurate virtual models of forest trees, for biomass measurement purposes. Both projects are funded by the University of Grenoble Alpes, through the AGIR framework. A PhD student, Romain Rombourg, is working on them.

9.4.1.2. Informal International Partners

The long term collaboration with TU Munich and Slobodan Ilic on human motion capture is ongoing with the work of Paul Huang [4] and [12] that was published at CVPR and IJCV this year. The work contributes with an approach that identifies and takes benefit of key poses when tracking shapes and 4D modeling.

9.5. International Research Visitors

9.5.1. Visits of International Scientists

9.5.1.1. Internships

Victoria Fernández Abrevaya

Date: 29th June 2015 - 27th September 2015

Institution: Universidad de Buenos Aires (Argentina)

Supervisor: Franck Hétroy-Wheeler

9.5.2. Visits to International Teams

9.5.2.1. Sabbatical programme

Reveret Lionel

Date: Jul 2014 - June 2015

Institution: Brown University (United States)

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific events organisation

10.1.1.1. General chair, scientific chair

- Edmond Boyer was general chair for 3DV 2015.

10.1.1.2. Member of the organizing committees

- Franck Hétroy-Wheeler was publication chair for 3DV 2015.
- Jean-Sébastien Franco was tutorial chair for 3DV 2015.

10.1.2. Scientific events selection

10.1.2.1. Chair of conference program committees

- Edmond Boyer was area chair for BMVC 2015.
- Jean-Sébastien Franco was area chair for 3DV 2015.

10.1.2.2. Member of the conference program committees

- Stefanie Wuhrer was on the program committee for: Eurographics Short Papers 2015, Eurographics Workshop 3DOR 2015, SOCG Multimedia Exposition 2015.
- Lionel Reveret was on the program committee for: Symposium on Computer Animation 2015.

10.1.2.3. Reviewer

- Edmond Boyer has reviewed for: CVPR 2015, ICCV 2015, CVMP 2015, SIGGRAPH 2015, ORASIS 2015.
- Jean-Sébastien Franco has reviewed for: CVPR 2015, ICCV 2015.
- Franck Hétroy-Wheeler has reviewed for: 3DV 2015.
- Lionel Reveret has reviewed for: SIGGRAPH 2015, SIGGRAPH Asia 2015, Symposium on Computer Animation 2015, Motion in Game 2015, EUROGRAPHICS 2016.
- Stefanie Wuhrer has reviewed for: ICCV 2015, SIGGRAPH Asia 2015.
- Julien Pansiot has reviewed for: MobiHealth 2015, ICCV 2015.

10.1.3. Journal

10.1.3.1. Member of the editorial boards

- Edmond Boyer is associate editor of IJCV.

10.1.3.2. Reviewer - Reviewing activities

- Edmond Boyer has reviewed for: IJCV, IEEE Transactions on PAMI.
- Lionel Reveret has reviewed for: TOG, Computer and Graphics.
- Stefanie Wuhrer has reviewed for: Graphical Models, IEEE Transactions on Visualization and Computer Graphics, The Visual Computer.
- Julien Pansiot has reviewed for: Journal of 3D Research.

10.1.4. Scientific expertise

- Edmond Boyer was on the evaluation committees of: European Research Council (evaluation of consolidator ERC proposals), HCERES (evaluation of the LP2N lab at Bordeaux) and ANR (French academic projects).
- Edmond Boyer was on the recruitment committees for CR2 positions at the Inria Grenoble Rhône-Alpes.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Licence: J.S. Franco, Algorithmics, 60h, Ensimag 1st year, Grenoble INP

Licence: J.S. Franco, C Project, 30h, Ensimag 1st year, Grenoble INP

Master: J.S. Franco, End of study project (PFE) Project Tutoring, 18h, Ensimag 2nd year, Grenoble INP

Master: J.S. Franco, 3D Graphics, 50h, Ensimag 2nd year, Grenoble INP

Master: J.S. Franco, Modelisation et programmation C++, 9h, Ensimag 2nd year, Grenoble INP

Master: J.S. Franco, Introduction to Computer Vision, 27h, Ensimag 1st year, Grenoble INP

Master: J.S. Franco, co-responsability of the Graphics, Vision, Robotics specialty of the Mosig Masters program, Second year Masters, Grenoble INP, Université Joseph Fourier

Licence: F. Hétroy-Wheeler, Algorithmics and programming, 45h, Ensimag dual education through apprenticeship 1st year, Grenoble INP

Master: F. Hétroy-Wheeler, Algorithmics and Discrete Optimisation, 18h, Ensimag 2nd year, Grenoble INP

Master: Edmond Boyer, 3D Modeling, 9h, M2R GVR, Université Joseph Fourier Grenoble, France.

Master: Edmond Boyer, projet de programmation, 30h, M1 informatique - M1 MoSig, Université Joseph Fourier Grenoble, France.

Master: Edmond Boyer, Introduction to Image Analysis, 15h, M1 MoSig, Université Joseph Fourier Grenoble, France.

Master: J. Pansiot, Introduction to Computational Anatomy, 2h, Grenoble University Hospital.

Master: J. Pansiot, Introduction to Computer Vision, 27h, Ensimag 3rd year, Grenoble INP.

10.2.2. Supervision

HdR: Franck Hétroy-Wheeler, Segmentation and skeleton methods for digital shape understanding, Université Grenoble Alpes, 20/11/2015.

PhD in progress : Benjamin Allain, Geometry and Appearance Analysis of Deformable 3D shapes, Université de Grenoble, started 01/10/2012, supervised by J.S. Franco and E. Boyer.

PhD in progress: Adnane Boukhayma, 4D model synthesis, Universiteé de Grenoble, started 01/10/2013, supervised by Edmond Boyer.

PhD in progress: Georges Nader, Evaluation de la qualité perceptuelle de maillages dynamiques et ses applications, Université Claude Bernard - Lyon 1, started 01/10/2013, supervised by Florent Dupont, Kai Wang and Franck Hétroy-Wheeler.

PhD in progress: Romain Rombourg, Digital tree: from the acquisition to a high-level geometric model, Université Grenoble Alpes, started 01/10/2015, supervised by Franck Hétroy-Wheeler and Eric Casella.

PhD in progress : Vagia Tsiminaki, Appearance Modelling and Time Refinement in 3D Videos, Université de Grenoble, started 01/10/2012, supervised by J.S. Franco and E. Boyer.

PhD in progress: Li Wang, Transport optimal pour l'analyse de formes en mouvement, Université de Grenoble, started 01/10/2013, supervised by Edmond Boyer and Franck Hétroy-Wheeler.

PhD in progress: Timo Bolkart, Statistical analysis of 3D human faces is motion, Saarland University, started 01/01/2012, supervised by Stefanie Wuhrer.

PhD in progress: Aurela Shehu, Geometric processing of near-isometrically deforming surfaces, Saarland University, started 01/04/2012, supervised by Stefanie Wuhrer.

PhD in progress : Jinlong Yang, Learning shape spaces of dressed 3D human models in motion, Université de Grenoble, started 01/10/2015, supervised by Franck Hétroy-Wheeler and Stefanie Wuhrer.

10.2.3. Juries

- Edmond Boyer was president of the Habilitation committee of Vincent Lepetit at Grenoble University.
- Jean-Sébastien Franco was examiner of the PhD thesis of Kun Liu (Université de Lorraine, Inria Nancy)
- Franck Hétroy-Wheeler is a member of the PhD monitoring committee of Van Tho Nguyen (University of Lorraine, INRA Nancy and Office National des Forêts).

10.3. Popularization

- Edmond Boyer gave invited talks at the ENS Cachan and at the Inria Paris, and the inaugural talk at the Kinovis platform inauguration.
- Franck Hétroy-Wheeler gave two talks on digital geometry to secondary school pupils and their families.
- Interactive Kinovis platform demonstration over 2 days for the "Fête de la Science" (within Morpheo: M. Heudre and J. Pansiot).

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] A. DJELOUAH. *Multi-view object segmentation*, Universite Joseph Fourier, March 2015, https://hal.inria.fr/tel-01155592

[2] F. HÉTROY-WHEELER. *Segmentation and Skeleton Methods for Digital Shape Understanding*, Université Grenoble Alpes, November 2015, Habilitation à diriger des recherches, https://hal.inria.fr/tel-01246006

Articles in International Peer-Reviewed Journals

[3] A. DJELOUAH, J.-S. FRANCO, E. BOYER, F. LE CLERC, P. PEREZ. *Sparse Multi-View Consistency for Object Segmentation*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", 2015, pp. 1 - 14 [*DOI :* 10.1109/TPAMI.2014.2385704], https://hal.inria.fr/hal-01115557

[4] C.-H. HUANG, C. CAGNIART, E. BOYER, S. ILIC. *A Bayesian Approach to Multi-view 4D Modeling*, in "International Journal on Computer Vision (IJCV)", June 2015, 21 p. [*DOI :* 10.1007/S11263-015-0832-Y], https://hal.inria.fr/hal-01211810

[5] G. NADER, K. WANG, F. HÉTROY-WHEELER, F. DUPONT. *Just Noticeable Distortion Profile for Flat-Shaded 3D Mesh Surfaces*, in "IEEE Transactions on Visualization and Computer Graphics", December 2015 [*DOI :* 10.1109/TVCG.2015.2507578], https://hal.archives-ouvertes.fr/hal-01242271

[6] M. ROUHANI, A. D. SAPPA, E. BOYER. *Implicit B-Spline Surface Reconstruction*, in "IEEE Transactions on Image Processing", January 2015, vol. 24, n^{o} 1, pp. 22 - 32 [*DOI :* 10.1109/TIP.2014.2366374], https://hal.inria.fr/hal-01101037

[7] L. WANG, F. HÉTROY-WHEELER, E. BOYER. *A Hierarchical Approach for Regular Centroidal Voronoi Tessellations*, in "Computer Graphics Forum", 2015, 15 p. , https://hal.inria.fr/hal-01185210

International Conferences with Proceedings

[8] B. ALLAIN, J.-S. FRANCO, E. BOYER. *An Efficient Volumetric Framework for Shape Tracking*, in "CVPR 2015 - IEEE International Conference on Computer Vision and Pattern Recognition", Boston, United States, June 2015, https://hal.inria.fr/hal-01141207

[9] T. BOLKART, S. WUHRER. *A Groupwise Multilinear Correspondence Optimization for 3D Faces*, in "IEEE International Conference on Computer Vision (ICCV)", Santiago, Chile, December 2015, The official version of this article is available on the IEEE website, https://hal.inria.fr/hal-01205460

[10] A. BOUKHAYMA, E. BOYER. *Video based Animation Synthesis with the Essential Graph*, in "3DV 2015 - International Conference on 3D Vision", Lyon, France, Proceedings of 3DV 2015, October 2015, https://hal.inria.fr/hal-01212168

[11] A. HEWER, I. STEINER, T. BOLKART, S. WUHRER, K. RICHMOND. *A statistical shape space model of the palate surface trained on 3D MRI scans of the vocal tract*, in "18th International Congress of Phonetic Sciences", Glasgow, United Kingdom, August 2015, https://hal.inria.fr/hal-01192790

[12] C.-H. HUANG, E. BOYER, B. DO CANTO ANGONESE, N. NAVAB, S. ILIC. *Toward User-specific Tracking by Detection of Human Shapes in Multi-Cameras*, in "CVPR 2015 - IEEE International Conference on Computer Vision and Pattern Recognition", Boston, United States, June 2015, https://hal.inria.fr/hal-01148449

Conferences without Proceedings

[13] G. NADER, K. WANG, F. HÉTROY, F. DUPONT. *Modèle Perceptuel pour la Détection du Seuil de Visibilité des Distorsions sur un Maillage 3D*, in "28èmes journées de l'Association Française en Informatique Graphique", Lyon, France, November 2015, https://hal.archives-ouvertes.fr/hal-01245373

Scientific Books (or Scientific Book chapters)

[14] A. BRUNTON, A. SALAZAR, T. BOLKART, S. WUHRER. *Statistical Shape Spaces for 3D Data: A Review*, in "Handbook of Pattern Recognition and Computer Vision 5th Edition", C. H. CHEN (editor), March 2016, https://hal.inria.fr/hal-01205998

Research Reports

[15] B. ALLAIN, L. WANG, J.-S. FRANCO, F. HETROY-WHEELER, E. BOYER. *Shape Animation with Combined Captured and Simulated Dynamics*, ArXiv, January 2016, n° arXiv:1601.01232, 11 p. , https://hal.inria.fr/hal-01255337

[16] M. DOUZE, J.-S. FRANCO, B. RAFFIN. *QuickCSG: Arbitrary and Faster Boolean Combinations of N Solids*, Inria - Research Centre Grenoble – Rhône-Alpes ; Inria, March 2015, n° RR-8687, https://hal.inria.fr/hal-01121419

[17] L. PISHCHULIN, S. WUHRER, T. HELTEN, C. THEOBALT, B. SCHIELE. *Building Statistical Shape Spaces for 3D Human Modeling*, Max-Planck Institute for Informatics, March 2015, https://hal.inria.fr/hal-01136221

Scientific Popularization

[18] E. BOYER, A. HILTON, C. LOSCOS. *Dynamic Geometry Reconstruction*, in "Digital Representations of the Real World: How to Capture, Model, and Render Visual Reality", M. MAGNOR, O. GRAU, O. SORKINE-HORNUNG, C. THEOBALT (editors), CRC Press, 2015, pp. 179-192, https://hal.inria.fr/hal-01159497