Year 2011

IRIT - VORTEX Team

- IMAGES & 3D
- MULTIMEDIA

CNRS - INPT - UPS - UT1 - UTM

ENSEEIHT

Presentation

The goal of this small book is to introduce visitors or potential partners to the research interest and competences developed in the part of the VORTEX team (Visual Object from Reality to EXpression) located on the ENSEEIHT site. The content of this booklet, for each topic, is very brief in order to keep it concise and short. We hope, however, that this material will help inspiring questions and interests from the readers and hopefully lead to fruitful collaborations that will be, in turn, also part of the future version of this booklet!

The booklet is organized into two parts, one on the **analysis of virtual and natural objects**: a first goal is to reconstruct an object from its view in one, two or a sequence of images. The reconstruction problem has been applied in different contexts: first from a single image of an object (e.g. shape-from-shading based on machine learning techniques, and bubble tag detection). Then stereo-vision techniques (from two, or a sequence of images) are studied and used to reconstruct quadrics or natural plants. A tracking algorithm to estimate the motion of a known point-based model within a video sequence is proposed. The second part concentrates on **accessing visual objects**. Strategies, models and compression schemes for streaming 2D satellite images and 3D content have been proposed. We have adapted images, videos and augmented reality content for mobile devices.

Team personal
December 1, 2011

Permanent:

- Vincent Charvillat, 05.34.32.21.71, charvi@n7.fr
- Jean Conter, 05.34.32.21.78, Conter@n7.fr
- Jean-Denis Durou, 05.34.32.21.72, durou@irit.fr
- Romulus Grigoras, 05.34.32.21.76, Romulus.Grigoras@n7.fr
- Pierre Gurdjos, 05.34.32.21.72, pgurdjos@n7.fr
- Zouhaier Hamrouni, 05.34.32.21.75, hamrouni@n7.fr
- Géraldine Morin, 05.34.32.21.70, morin@n7.fr
- Philippe Puech, 05.34.32.21.77, Puech@n7.fr
- Bernard Thiesse, 05.34.32.21.74, thiesse@n7.fr

Non-permanent:

- Benjamin Ahsan, benjamin.ahsan@n7.fr
- Benoît Baccot, Benoit.Baccot@irit.fr
- Lilian Calvet, lilian.calvet@n7.fr
- Axel Carlier, axel.carlier@n7.fr
- Quoc-Viet Dang, quocviet.dang@n7.fr
- Jérôme Guénard, jguenard@n7.fr
- Pauline Julian, p.julian@n7.fr
- Thi Phuong Nghiem, thiphuong.nghiem@n7.fr
- Alina Nitu, alina-nitu88@yahoo.com
- Viorica Pătrăucean, vpatrauc@n7.fr
- Christophe Piombo, christophe.piombo@n7.fr
- Arash Shafiei, arash.shafiei@gmail.com

Is Dual Linear Self-Calibration Artificially Ambiguous?

Keywords: Structure-from-motion, Self-calibration, Closed-form Solution, Dual Absolute Quadric, Critical Motions, Confocal Quadrics.

This purely theoretical work investigates the problem of artificial singularities in camera self-calibration. Self-calibration allows one to upgrade a projective reconstruction to metric and has a concise and well-understood formulation based on the Dual Absolute Quadric (DAQ), a rank-3 quadric envelope satisfying (nonlinear) spectral constraints: it must be positive of rank 3. The practical scenario we consider is the one of square pixels, known principal point and varying unknown focal length, for which generic Critical Motion Sequences (CMS) have been thoroughly derived. The standard linear self-calibration algorithm uses the DAQ paradigm but ignores the spectral constraints. It thus has artificial CMSs, which have barely been studied so far. We propose an algebraic model of singularities based on the confocal quadric theory. It allows to easily derive all types of CMSs. We first review the already known generic CMSs, for which any self-calibration algorithm fails. We then describe all CMSs for the standard linear self-calibration algorithm; among those are artificial CMSs caused by the above spectral constraints being neglected. We then show how to detect CMSs. If this is the case it is actually possible to uniquely identify the correct self-calibration solution, based on a notion of signature of quadrics. The main conclusion of this paper is that a posteriori enforcing the spectral constraints in linear self-calibration is discriminant enough to resolve all artificial CMSs.

VORTEX participant: Pierre Gurdjos.

External participants: Adrien Bartoli (LASMEA, Clermont-Ferrand), Peter Sturm (INRIA Grenoble Rhône-Alpes).

Associated publication:

Pierre Gurdjos, Peter Sturm, Adrien Bartoli, *Is Dual Linear Self-Calibration Artificially Ambiguous?*, 12$^{\text{th}}$ IEEE International Conference on Computer Vision (ICCV 2009), Tokyo, Japan, September-October 2009.

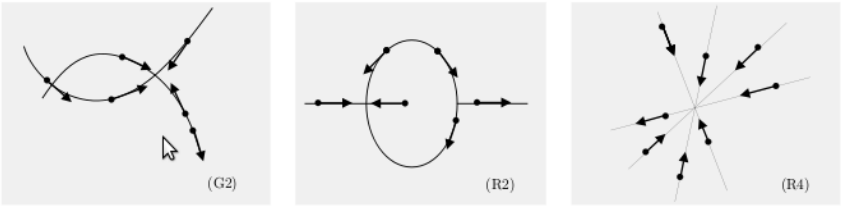


Illustration 1 - The three classes (G2), (R2) and (R4) of artificial CMS completing the known generic CMS. The ambiguity for the DAQ is 1D if the CMS is rank-2 i.e., it includes at least three cameras with at least one on the circle for (R2).

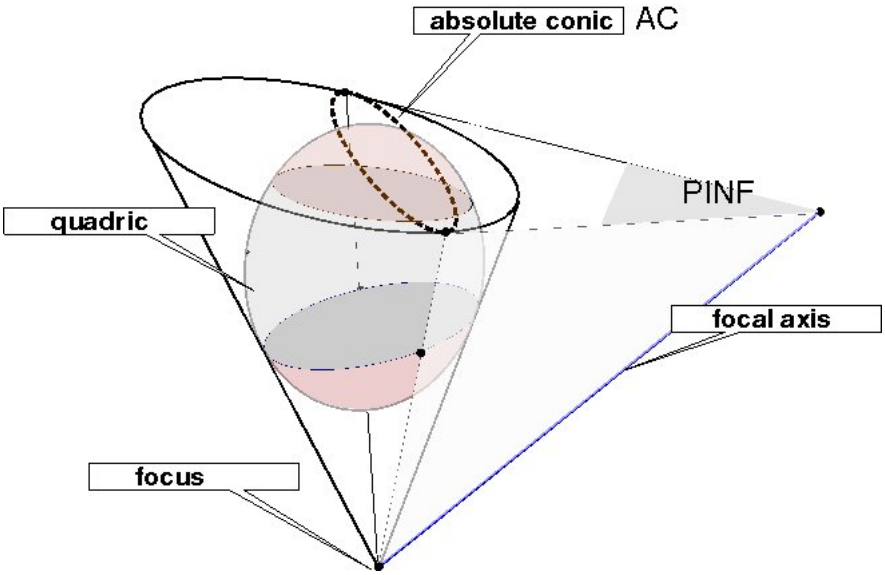


Illustration 2 - A focus and a focal axis of a quadric respectively coincide with the vertex and the revolution axis of a circular cone (i.e., having double contact with Ω_∞) enveloping the quadric.

Multiple View Reconstruction of a Quadric of Revolution from its Occluding Contours

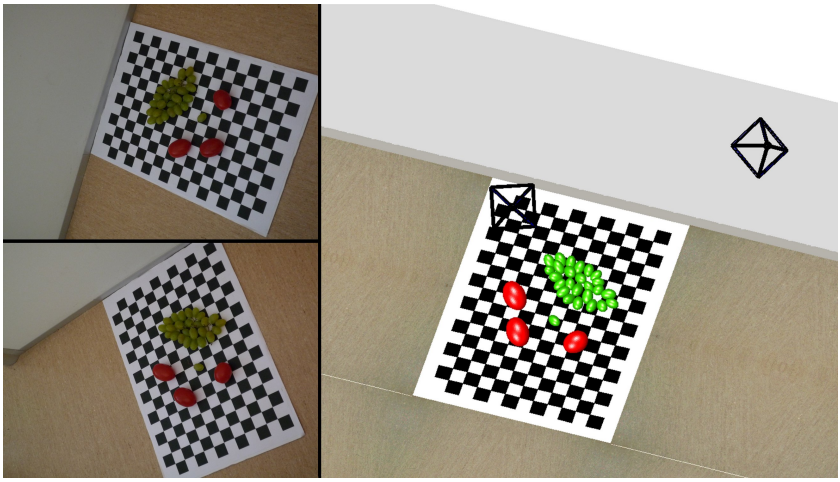
Keywords: 3D Reconstruction, Quadrics of Revolution.

The problem of reconstructing a quadric from its occluding contours is one of the earliest problems in computer vision. It is known that three contours from three views are required for this problem to be well-posed while Cross et al. have proved in that, with only two contours, what can be obtained is a 1D linear family of solutions in the dual projective space. In this work, we describe a multiple view algorithm that unambiguously reconstructs so-called Prolate Quadrics of Revolution (PQoRs, see text), given at least two finite projective cameras. In particular, we show how to obtain a closed-form solution. The key result on which is based this work is a dual parameterization of a PQoR, using a 7-dof linear combination of the quadric dual to the principal focus pair and the Dual Absolute Quadric (DAQ). One of the contributions is to prove that the images of the principal foci of a PQoR can be recovered set-wise from the images of the PQoR and the DAQ. The performance of the proposed algorithm is illustrated on simulations and experiments with real images.

VORTEX participants: Vincent Charvillat, Jérôme Guénard, Pierre Gurdjos, Géraldine Morin.

Associated publication:

Pierre Gurdjos, Jérôme Guénard, Vincent Charvillat, Géraldine Morin, *Multiple View Reconstruction of a Quadric of Revolution from its Occluding Contours*, 9$^{\text{th}}$ Asian Conference on Computer Vision (ACCV 2009), Xi'an, China, September 2009.



Quadrics of revolution - Left: two images of the scene. Right: quadrics and camera reconstruction.

Reconstruction of Virtual Vine Models from Images

Keywords: Plant Modelisation, Analysis-by-synthesis, Image Processing, 3D Reconstruction.

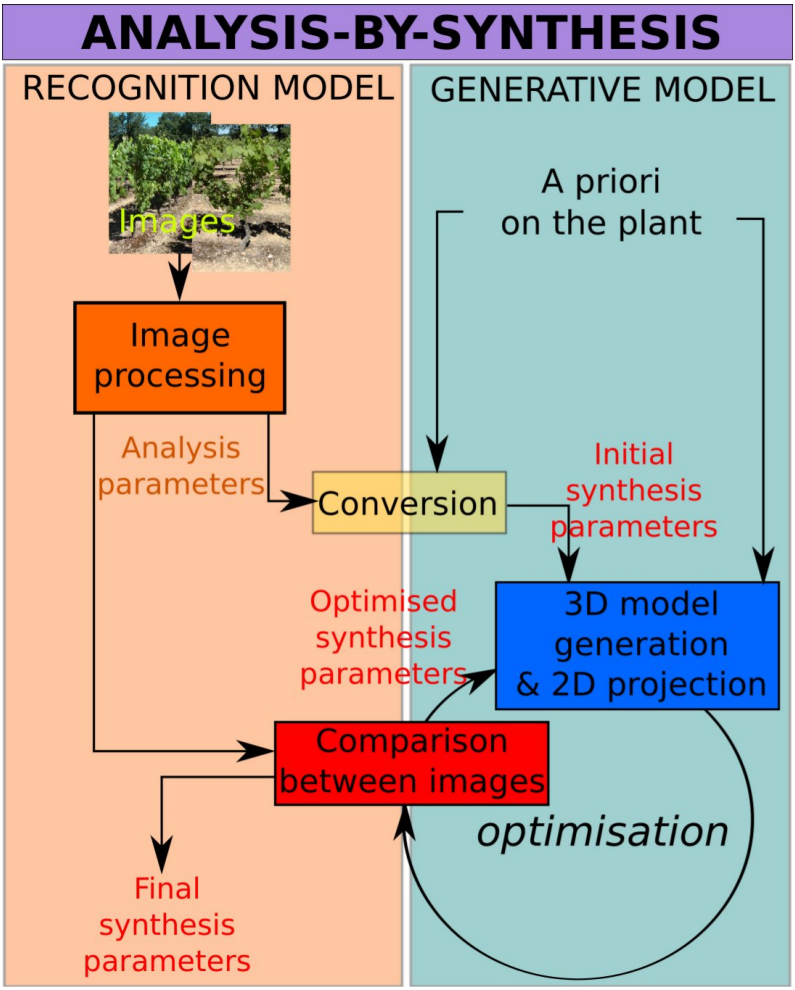
We propose an analysis by synthesis method for modeling plants from images. The purpose is to understand an image and explain it providing a number of a priori knowledge about the 3D scene. Initially, we get an approximation of the plant modeled in 3D. Then, comparing its reprojection with the original image, we refine this model through an iterative optimisation process. We work on images of vines. Our goal is not to represent the plant positioning the leaves one by one but to obtain a realistic model, consistent with the images.

VORTEX participants: Vincent Charvillat, Jérôme Guénard, Pierre Gurdjos, Géraldine Morin.

External participant: Frédéric Boudon (Virtual Plants team, INRIA Montpellier).

Associated publication:

Jérôme Guénard, Géraldine Morin, Frédéric Boudon, Pierre Gurdjos, Vincent Charvillat, *Reconstruction de modèles virtuels de vignes à partir d'images*, ORASIS 2011, Praz-sur-Arly, June 2011.



Our approach - On the left side, the recognition model explains how we get information from the images for the 3D reconstruction. On the right side, the generative model explains how use a priori knowledge about the scene for the 3D reconstruction.

Camera Tracking from Plans Markers

Keywords: Bundle Adjustment, Camera Tracking, Concentric Circles, Fiducial Marker.

In this paper we present a marker-based camera tracking with fiducial markers made up of a set of concentric circles. A set of concentric circles (C^2 Tag) is well suited to cope with the camera tracking problem: a C^2 Tag is very discriminatory in pictures, its detection is also possible in occluded cases and it can be identified by mean of a possible coding. In addition, its geometry is closely related to the Euclidean structure of the support plane. We present the general paradigm of camera tracking based on two coplanar C^2 Tags and the different steps involved: detection, identification, adjustment, 2D reconstruction, calibration and 3D reconstruction. Our contribution focuses on the last step which consists of a bundle adjustment: we describe a minimal parametrization and a criterion which minimizes the distance between measured points and the ellipses constituting the marker. Experiments on synthetic data and real data confirm the performance of the proposed method.

VORTEX participants: Lilian Calvet, Vincent Charvillat, Pierre Gurdjos.

External participants: Simone Gasparini, Peter Sturm (INRIA Grenoble Rhône-Alpes).

Associated publication:

Lilian Calvet, Vincent Charvillat, Pierre Gurdjos, Simone Gasparini, Peter Sturm, *Suivi de caméra à partir de marqueurs plans composés de cercles concentriques: paradigme et algorithme*, ORASIS 2011, Praz-sur-Arly, June 2011.

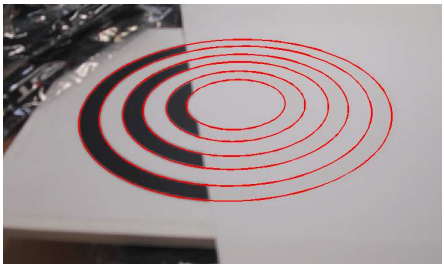
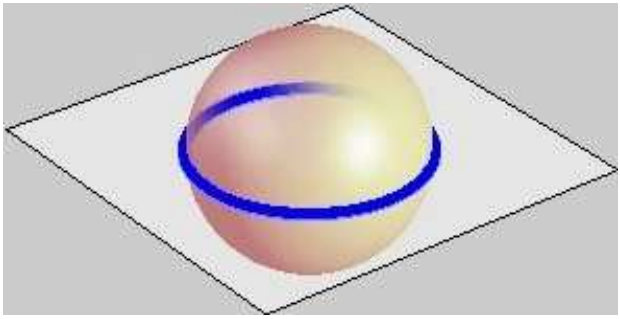
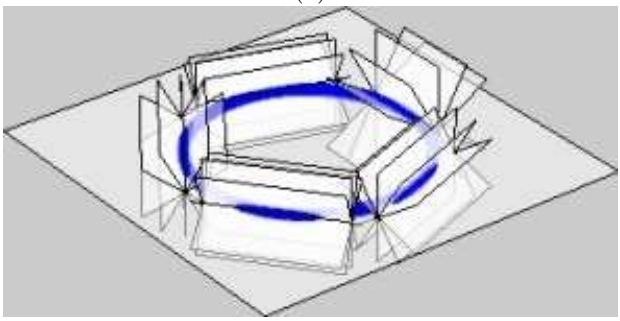


Illustration 1 - Detection of an occulted C^2 Tag. Detected ellipses appear in red.



(a)



(b)

Illustration 2 - (a) In the 3D projective space, a circle is a points locus, coresponding to the intersection between a sphere and a plane. (b) In the dual space, a circle is a degenerated quadric i.e., a 3D plans envelope.

Resolution of the Shape-from-Shading Problem using Machine Learning

Keywords: 3D Reconstruction, Shape-from-shading, Machine Learning.

In this work, we try to solve the shape-from-shading problem using machine learning techniques. A first object whose shape is known is imaged under a known lighting. Its photograph is cut into patches $\mathcal{I}_a$ of size 3×3, associated with normals $\mathbf{N}$. We process a principal components analysis (PCA) on the data, in order to project the patches $\mathcal{I}_a$ in a space of lower dimension n. Then, we analyse the photograph of a second object of unknown shape, taken under the same lighting, which we cut into patches $\mathcal{I}_t$ of the same size 3×3. For each patch $\mathcal{I}_t$, we search its closest patch $\mathcal{I}_a$, according to some distance in the space of dimension n, and assign to $\mathcal{I}_t$ the normal $\mathbf{N}$ associated to $\mathcal{I}_a$.

This method is tested on synthetic, as well as on real images. We show its advantages, but also its drawbacks, particularly faced to the well-known concave/convex ambiguity. Nevertheless, this work opens some perspectives for the resolution of the photometric stereo problem.

VORTEX participants: Vincent Charvillat, Maxime Daramy, Jean-Denis Durou, Pierre Gurdjos.

Associated publication:

Jean-Denis Durou, Vincent Charvillat, Pierre Gurdjos, Maxime Daramy, *Résolution du shape-from-shading par apprentissage*, ORASIS 2011, Praz-sur-Arly, June 2011.


Data - Left: Learning image. Right: Test image.


Results - Two images simulated using the normal field estimated, via PCA, from the images above. On the left, the initial lighting is re-used. On the right, another lighting is used, which shows that the estimated normal field is non-integrable.

Bubble Tag Identification

Keywords: Bubble Tags, Euclidean Reconstruction, Projective Invariant Signatures, Information Retrieval.

We have at our disposal a large database containing images of various configurations of coplanar circles, randomly laid-out, called "Bubble Tags". The images are taken from different viewpoints. Given a new image (query image), the goal is to find in the database the image containing the same bubble tag as the query image. We propose representing the images through projective invariant signatures which allow identifying the bubble tag without passing through an Euclidean reconstruction step. This is justified by the size of the database, which imposes the use of queries in vectorial form, i.e. not in matrix form. The experiments carried out confirm the efficiency of our approach, in terms of precision and complexity.

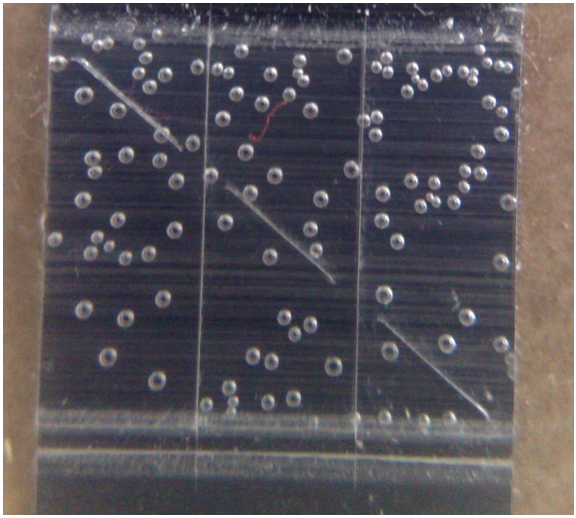
One of the applications of this work has been developped within the Geowine project, in collaboration with the AGRIMIP Innovation Pole.

This project aims at implementing a system for wine tracability and authentification. These objectives are to be achieved using Bubble Tag^{TM}, the solution proposed by Prooftag. The Bubble Tag^{TM} is the result of a chaotic self-generation of bubbles within a transparent polymer. Our tasks include Bubble Tag^{TM} lecture, signature extraction and identification among a large database, using our work presented above.

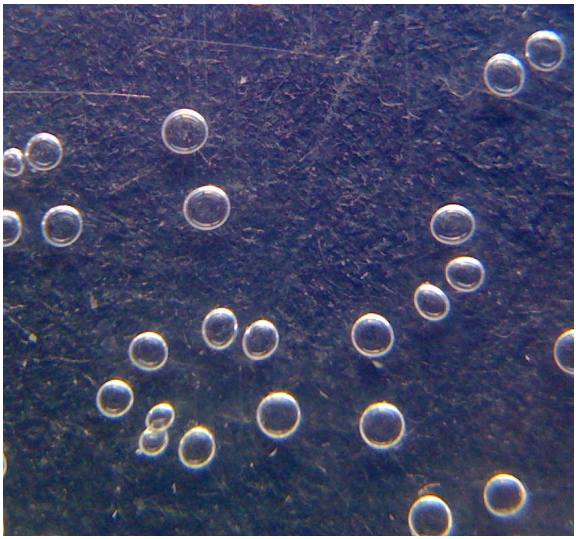
VORTEX participants: Jean Conter, Pierre Gurdjos, Viorica Pătrăucean.

Associated publication:

Viorica Pătrăucean, Pierre Gurdjos, Jean Conter, *Bubble Tag Identification Using an Invariant-Under-Perspective Signature*, 20th IAPR International Conference on Pattern Recognition (ICPR 2010), pages 408-411, Istanbul, Turkey, August 2010.



(a)



(b)

Example - An example of a Bubble Tag^{TM}. (a) The complete tag. (b) Zoom on bubbles. Note that the bubbles are not necessarily round and may overlap.

Point-based 3D Tracking

Keywords: Visual Tracking, Point-based Model, Splatting, GPGPU.

We use point-based 3D models for real-time visual object tracking with a single monocular camera. Previously, this problem as been addressed using sparse 3D models based on edges, meshes or textured patches. Instead we use a point-based model and related methods recently developed in the computer graphics community. The points are arbitrarily sampled on the object surface and no connectivity information is required. We show that state-of-the-art techniques for real-time rendering of point-based geometries can be efficiently repurposed for use in a 3D tracking context. We derive an original tracking algorithm from the method proposed earlier by Vacchetti et al. which combines an iterative pose update and a keyframe-based 3D registration. We propose a mathematically sound framework for using point-based models for the purpose of visual tracking. This framework allows the reconstruction of dense linear motion predictors and the generation of novel views from keyframes for wide baseline feature matching. We make use of the same general surface splatting technique, which we implement, together with other low-level vision tasks, on the GPU, leading to a real-time tracking algorithm.

Further results and video material illustrating the effectiveness of the system on live video sequences can be found at:
`http://dehais.perso.enseeiht.fr/pbmtracking`

VORTEX participants: Vincent Charvillat, Géraldine Morin.

External participant: Christophe Dehais (Fitting Box company).

Associated publication:

Vincent Charvillat, Géraldine Morin, Christophe Dehais, *From rendering to tracking point-based 3D models*, Image and Vision Computing 28(9):1386-1395, September 2010.

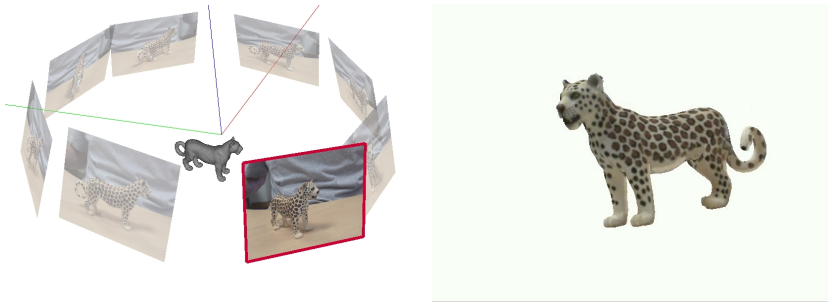


Illustration 1 - Left: A keyframe "cloud". Right: A novel view generated from one of the keyframes of figure on the left.

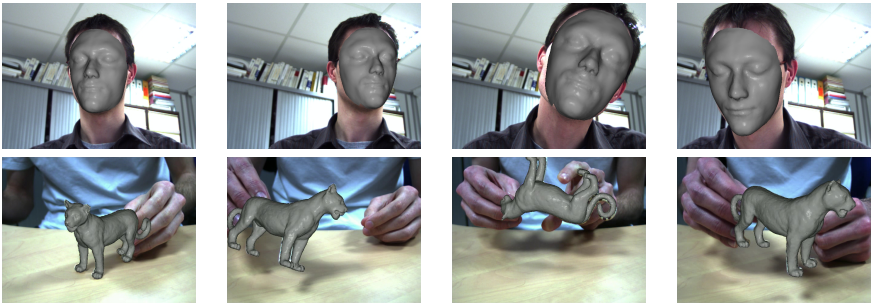


Illustration 2 - Screenshots from live tracking sequences. First line: tracking the toy leopard shown throughout the paper. Second line: Tracking using a face model aquired with a range scanner.

Automatic Hair Detection in the Wild

Keywords: Hair Segmentation, Shape Prior, Active Contours.

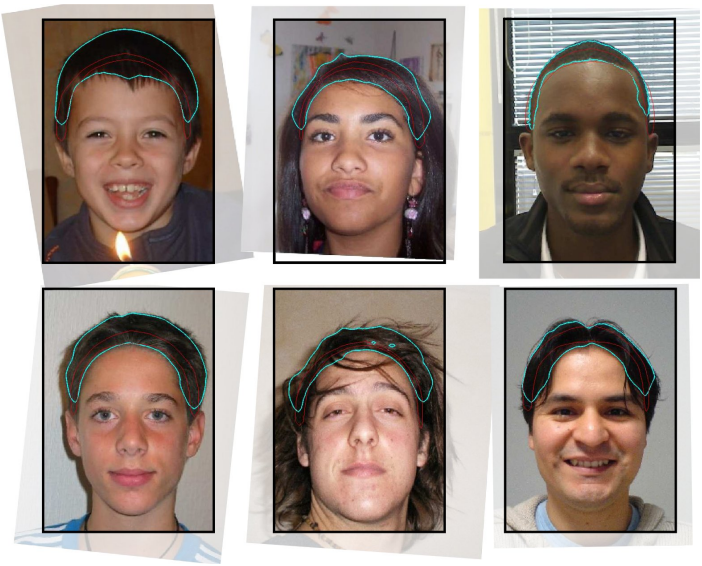
This paper presents an algorithm for segmenting the hair region in uncontrolled, real life conditions images. Our method is based on a simple statistical hair shape model representing the upper hair part. We detect this region by minimizing an energy which uses active shape and active contour. The upper hair region then allows us to learn the hair appearance parameters (color and texture) for the considered image. Finally, those parameters drive a pixel-wise segmentation technique that yields the desired (complete) hair region. We demonstrate the applicability of our method on several real images.

VORTEX participants: Vincent Charvillat, Pauline Julian.

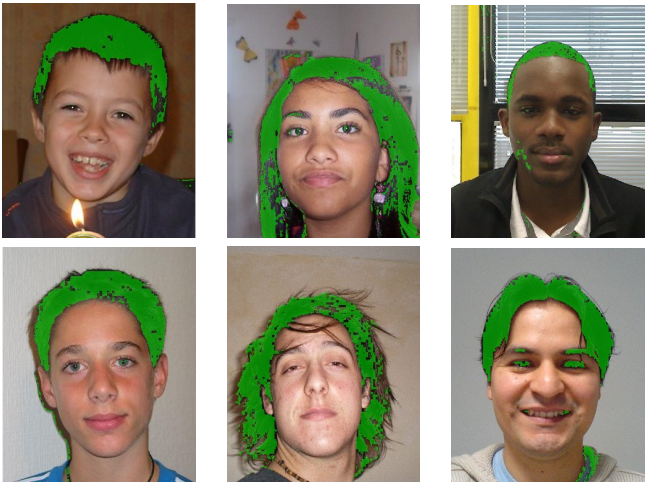
External participants: Adrien Bartoli (LASMEA, Clermont-Ferrand), Ariel Choukroun, Christophe Dehais (Fitting Box company), François Lauze (DIKU, Copenhagen, Denmark).

Associated publication:

Pauline Julian, François Lauze, Christophe Dehais, Vincent Charvillat, Ariel Choukroun, Adrien Bartoli, *Automatic Hair Detection in the Wild*, 20$^{\text{th}}$ IAPR International Conference on Pattern Recognition (ICPR 2010), pages 4617-4620, Istanbul, Turkey, August 2010.



Results 1 - USHM (Upper Hair Shape Model) results.



Results 2 - Pixel wise segmentation results.

Interactive Augmentation of Photographs Depicting Prehistoric Engravings

Keywords: Augmented Reality, Interactive Board, Prehistoric Engravings.

We devised a complete setup that eases the access to prehistoric features in an interactive fashion. High resolution photographs of engraved walls are combined with manual drawings made by experts. To make the interaction very intuitive, images of engraved panels are projected onto an electro-magnetic tracking board that is able to report the position and state of up to three electronic stylus. The user moves a stylus to reveal augmentations interactively by pointing directly the image. At a software level, several problems had to be overcome, from precisely registering line drawings of the engravings made manually by an expert onto the corresponding photographs, to the design of a usable and fault tolerant software. A complete system has been realized and successfully tested with a relatively wide audience, composed in particular of children. This work demonstrates the interest and effectiveness of 2D Augmented Reality in the context of cultural heritage.

As an extension we also created a mobile version of the AR application: the so-called MARCH application. MARCH stands for Mobile Augmented Reality for Cultural Heritage. Simply point the camera of a mobile device at prehistoric cave engravings. Then MARCH augments the captured images with the expert's drawings, highlighting in real time the animal engravings, which are almost impossible to observe with the naked eye. We have created a mobile augmented reality application which runs at 14 FPS for 320×240 frames on a Nokia N95 smartphone. We describe the optimizations and the requirements needed to obtain these results on mobile devices.

VORTEX participants: Vincent Charvillat, Omar Choudary, Jean Conter, Romulus Grigoras.

External participant: Christophe Dehais (Fitting Box company).

Associated publication:

Omar Choudary, Vincent Charvillat, Romulus Grigoras, Pierre Gurdjos, *MARCH: Mobile Augmented Reality for Cultural Heritage*, 17$^{\text{th}}$ ACM International Conference on Multimedia, pages 1023-1024, Beijing, China, October 2009.

Result - Apparition of a carving, thanks to our device.

3D Streaming: Streaming Progressive Mesh Models

Keywords: Streaming, 3D Data, Multiresolution, Progressive Meshes, Packetization, Networked Virtual Environment.

3D triangular meshes are becoming an increasingly prevalent data type in networked applications such as digital museums, online games, and virtual worlds. In these applications, a 3D mesh is typically coded progressively, yielding a multiresolution representation suitable for streaming. While such progressive coding allows incremental rendering for users while data is being transmitted, it introduces dependencies between data, causing delay in rendering when packets are lost. This article quantitatively analyzes the effects of such dependency by modeling the distribution of decoding time as a function of mesh properties and network parameters. We apply our model to study two extreme cases of dependency in progressive meshes and show that the effect of dependencies on decoded mesh quality diminishes with time. Our model provides the expected decoded mesh qualityat the receiver at a given time. Based on this expected value, we propose a packetization strategy that improves the decoded mesh quality during the initial stage of streaming. We validate the accuracy of our model under a variety of network conditions, including bursty losses, fluctuating RTT, and varying sending rate. The values predicted from our model match the measured value reasonably well in all cases except when losses are too bursty.

VORTEX participants: Romulus Grigoras, Géraldine Morin.

External participants: Wei Cheng, Wei Tsang Ooi (National University of Singapore), Sébastien Mondet (University of Oslo, Norway).

Associated publication:

Wei Cheng, Wei Tsang Ooi, Sébastien Mondet, Romulus Grigoras, Géraldine Morin, *Modeling Progressive Mesh Streaming: Does Data Dependency Matter?*, ACM Transactions on Multimedia Computing, Communications, and Applications 7(2), February 2011.

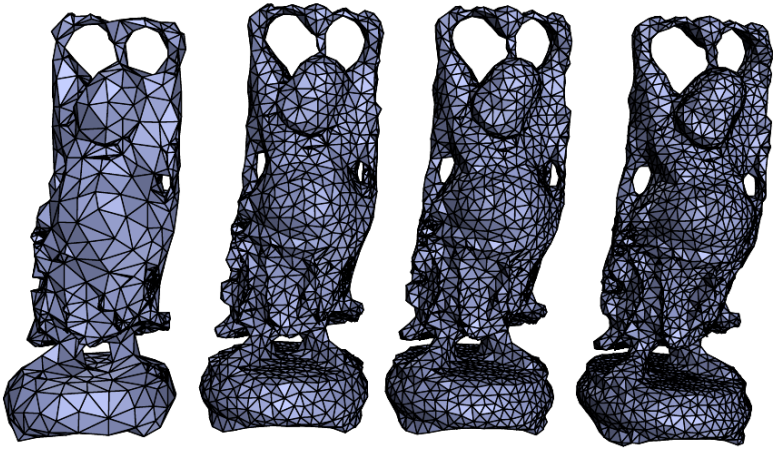


Illustration - Progressive codage of a 3D object.

3D Streaming: Plants streaming

Keywords: Streaming, Plant Models, Multiresolution, Progressive Coding, Progressive Transmission, Networked Virtual Environment.

Just as in the real world, plants are important objects in virtual worlds for creating pleasant and realistic environments, especially those involving natural scenes. As such, much effort has been made in realistic modeling of plants. As the trend moves towards networked and distributed virtual environments, however, the current models are inadequate as they are not designed for progressive transmissions. In this article, we fill in this gap by proposing a progressive representation for plants based on generalized cylinders. We model the shape and thickness of branches in a plant as Bézier curves, group the curves according to the similarity, and differentially code the curves to represent the plant in a compact and progressive manner. To facilitate the transmission of the plants, we quantify the visual contribution of each branch and use this weight in packet scheduling. We show the efficiency of our representations and the effectiveness of our packet scheduler through experiments over a wide area network.

VORTEX participants: Romulus Grigoras, Géraldine Morin.

External participants: Frédéric Boudon (Virtual Plants team, INRIA Montpellier), Wei Cheng, Wei Tsang Ooi (National University of Singapore), Sébastien Mondet (University of Oslo, Norway).

Associated publication:

Sébastien Mondet, Wei Cheng, Géraldine Morin, Romulus Grigoras, Frédéric Boudon, Wei Tsang Ooi, *Compact and Progressive Plants Models for Streaming in Networked Virtual Environments*, ACM Transactions on Multimedia Computing, Communications, and Applications 5(3), August 2009.

Illustration - Frontpage of the journal "ACM Transactions on Multimedia Computing, Communications, and Applications" in August 2009.

Satellite Image Compression by Post-Transforms in the Wavelet Domain

Keywords: Satellite, Still Image Coding, Discrete Transforms, Wavelet Transform.

The need for efficient compression schemes on board Earth observation satellites increases as the image size augments and the resolution enhances. This paper proposes a new compression scheme adapted to on-board constraints. The proposed coding scheme associates the wavelet transform to a post-transform to improve decorrelation and thus compression performance. Experiments on Earth observation images have already shown significant gain over the simple wavelet transform. Then, an appropriate progressive coder can take advantage of the post-transform energy compaction. Finally, this method should compete state-of-the-art compression schemes in terms of rate-distortion trade-off and targets simple hardware implementation on board spacecrafts.

VORTEX participants: Vincent Charvillat, Géraldine Morin.

External participants: Marie Chabert (IRIT-SC team), Xavier Delaunay (Noveltis company).

Associated publication:

Xavier Delaunay, Marie Chabert, Vincent Charvillat, Géraldine Morin, *Satellite Image Compression by Post-Transforms in the Wavelet Domain*, Signal Processing 90(2):599-610, February 2010.

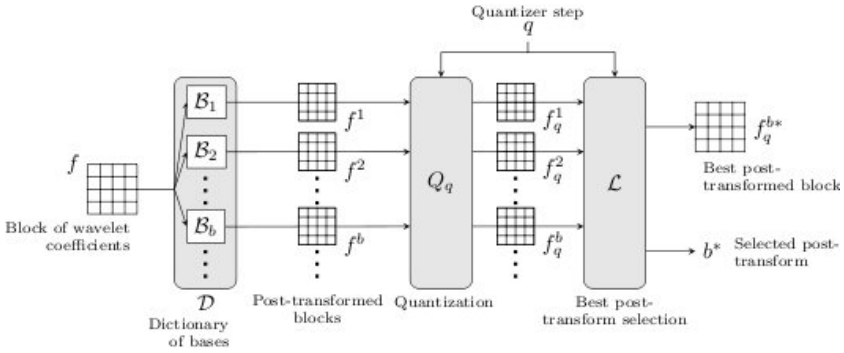


Illustration 1 - Post-transform compression scheme.

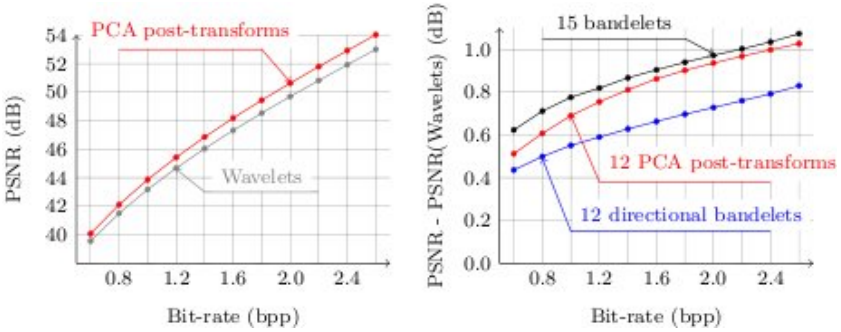


Illustration 2 - Mean compression results on six Earth observation images. Post-transforms with 12 PCA bases is compared to bandelet transforms with the 12 directional bases and to bandelet transform with 15 bases i.e. with the DCT and the two Haar bases added.

Crowdsourced Automatic Zoom and Scroll for Video Retargeting

Keywords: Crowd-sourcing, Video Retargeting, Automatic Zoom and Pan, Zoomable Video.

Screen size and display resolution limit the experience of watching videos on mobile devices. The viewing experience can be improved by determining important or interesting regions within the video (called regions of interest, or ROIs) and displaying only the ROIs to the viewer. Previous work focuses on analyzing the video content using visual attention model to infer the ROIs. Such content-based technique, however, has limitations. In this paper, we propose an alternative paradigm to infer ROIs from a video. We crowdsource from a large number of users through their implicit viewing behavior using a zoom and pan interface, and infer the ROIs from their collective wisdom. A retargeted video, consisting of relevant shots determined from historical users behavior, can be automatically generated and replayed to subsequent users who would prefer a less interactive viewing experience. This paper presents how we collect the user traces, infer the ROIs and their dynamics, group the ROIs into shots, and automatically reframe those shots to improve the aesthetics of the video. A user study with 48 participants shows that our automatically retargeted video is of comparable quality to one handcrafted by an expert user.

VORTEX participants: Axel Carlier, Vincent Charvillat, Romulus Grigoras, Géraldine Morin.

External participant: Wei Tsang Ooi (National University of Singapore).

Associated publication:

Axel Carlier, Vincent Charvillat, Romulus Grigoras, Géraldine Morin, Wei Tsang Ooi, *Crowdsourced Automatic Zoom and Scroll for Video Retargeting*, 18th ACM International Conference on Multimedia, pages 201-210, Florence, Italy, October 2010.

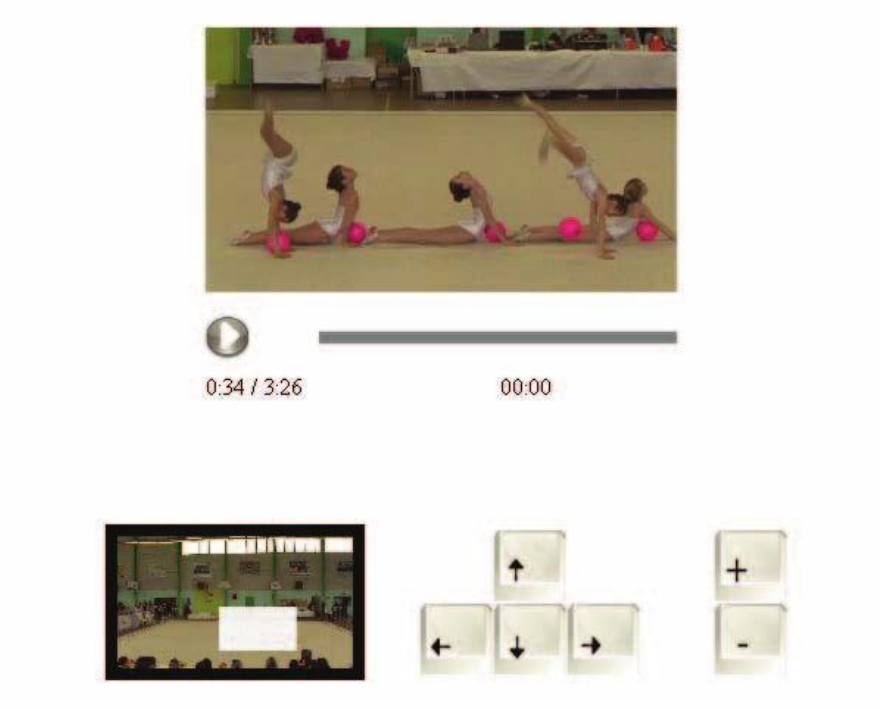


Illustration - User interface of the zoomable video player.

Combining Content-based Analysis and Crowdsourcing

Keywords: Crowd-sourcing, Content-Analysis, Zoomable Video, Interaction Techniques.

This paper introduces a new paradigm for interacting with zoomable video. Our interaction technique reduces the number of zooms and pans required by providing recommended viewports to the users, and replaces multiple zoom and pan actions with a simple click on the recommended viewport. The efficacy of our technique lies on the quality of the recommended viewport, which needs to match the user intention, track movement in the scene, and frame the scene in the video properly. To this end, we propose a hybrid method where content analysis and crowdsourcing are used to complement each other to recommend the viewports. We first compute a preliminary sets of recommended viewports by analyzing the content of the video. These viewports allow tracking of moving objects in the scene, and are framed without violating basic aesthetic rules. To improve the relevance of the recommended viewport, we collect viewing statistics as users view a video, and use the viewports they select to reinforce the importance of certain recommendations and penalize others. New recommendations that are not previously recognized by content analysis may also emerge. The resulting recommended viewports converge towards the regions in the video that are relevant to users. A user study involving 70 participants shows that a user interface build with our paradigm leads to more zooms into more informative regions with fewer interactions required.

VORTEX participants: Axel Carlier, Vincent Charvillat.

External participants: Ravindra Guntur, Wei Tsang Ooi (National University of Singapore).

Associated publication:

Axel Carlier, Ravindra Guntur, Vincent Charvillat, Wei Tsang Ooi, *Combining Content-based Analysis and Crowdsourcing to Improve User Interaction with Zoomable Video*, 19th ACM International Conference on Multimedia, pages 43-52, Scottsdale, USA, November/December 2011.

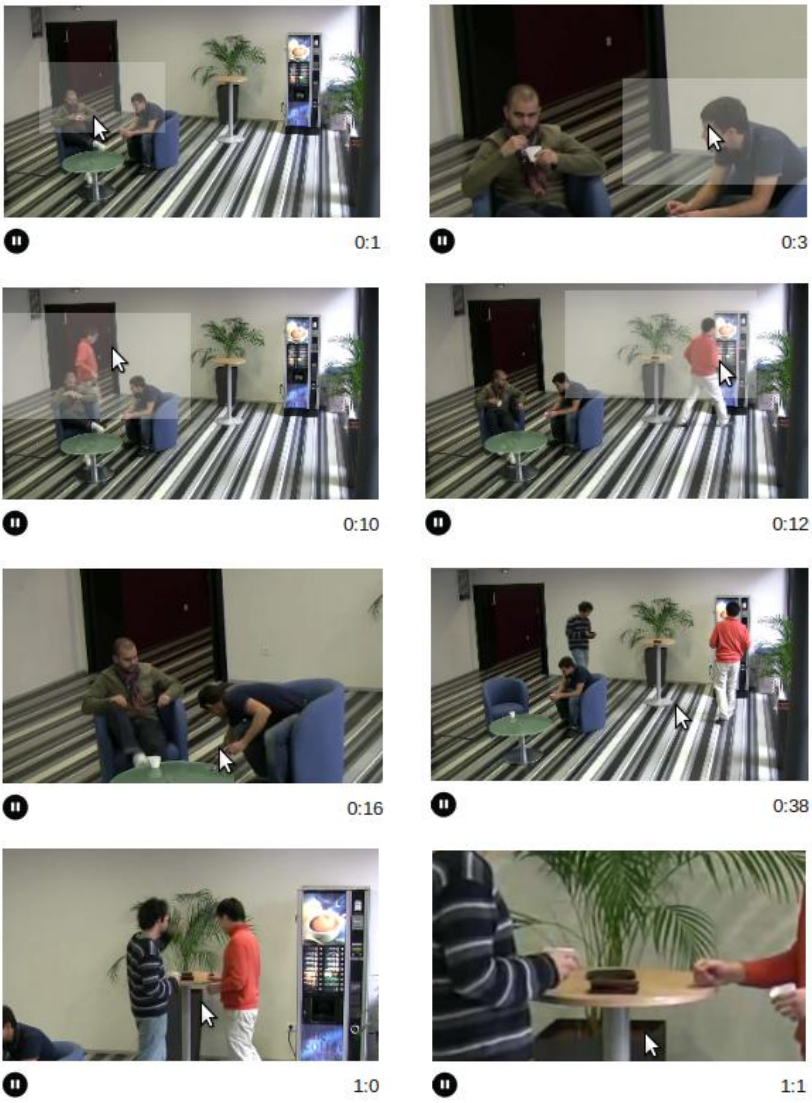


Illustration - Our Zoomable Interface.

Reinforcement Learning for Online Optimization of Banner Format and Delivery

Keywords: Rich Media Advertiving, Banner Sequences, Advertisement Impact, Reinforcement Learning, Online Advertising, Stochastic Models.

In our Internet-connected world, online advertising has grown into one of the most successful advertising channels, since users spend an important amount of time browsing the web. Among the different types of online advertising (emails, games, etc.), we are particularly interested in contextual ads using rich media banners that display motion and exploit sensory information such as video, audio, animation etc. Once the various banners of an ad campaign are produced, a legitimate question arises for a web marketer: among various options, and for the same banner content, what is the optimal banner format and delivery policy?

In this chapter, we deal with three main problems a web marketer can be confronted with: the right format among those available (e.g. text, image, video, interactive, etc.), the right time to display a banner during a user browsing session (e.g. at the beginning, at the end or when salient events are detected, etc.) and the right sequence of banners to display (that takes into account the format and the time problem). We show that these problems share common points. These points fit well within the reinforcement learning framework: a "trial-and-error" process can be used to dynamically determine an advertising policy that optimizes a criterion based on an impact measure (e.g. the click-through rate or the session duration of a user). Two stochastic models based on Markov Decisional Processes and Multi-Armed Problems are presented in order to solve the three problems.

Results, showing the power and the efficiency of the two models to solve our problems, are also given. By comparing to a "ground truth" acquired by observing user browsing session on a test site, we conclude that our models are able to determine optimal advertising policies concerning banner formats and delivery.

VORTEX participants: Benoît Baccot, Vincent Charvillat, Omar Chouday, Romulus Grigoras.

Associated publication:

Benoît Baccot, Omar Chouday, Romulus Grigoras, Vincent Charvillat, *On the Impact of Sequence and Time in Rich Media Advertising*, 17th ACM International Conference on Multimedia, pages 849-852, Beijing, China, October 2009.

Illustration 1 - The test website and the three types of banners.

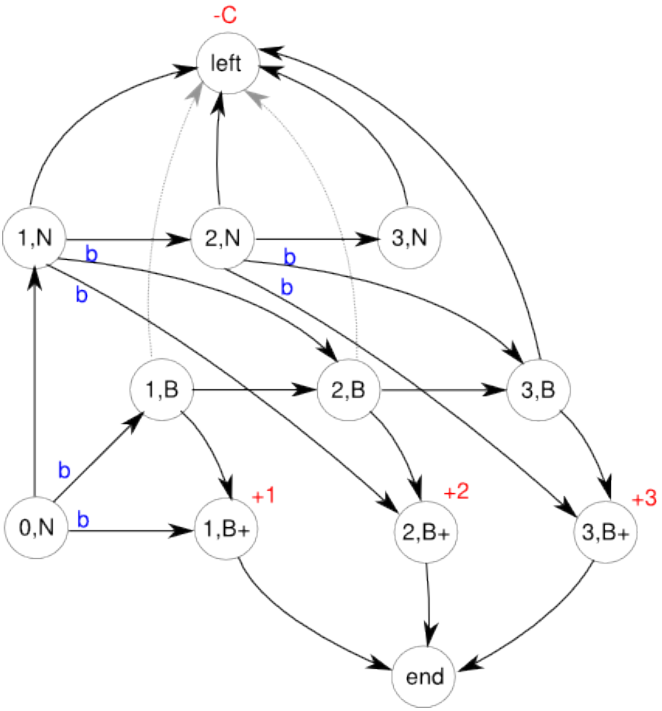


Illustration 2 - A decisional model used to choose the best moment to display a banner.

Towards Peer-Assisted Rendering in Networked Virtual Environments

Keywords: Peer-Assisted Rendering, Image-based Rendering, Networked Virtual Environments.

This paper introduces a new technique, called peer-assisted rendering, that aims to enable interactive navigation in a 3D networked virtual environment using a resource-constrainst device, by speeding up the rendering. A resource-constrained client requests part of the rendered scenes from other peers with similar viewpoints within the virtual environment, and merges the rendered parts into its own view. This approach is more scalable than the previous solutions based on server-based pre-rendering. The goal of this paper is to make a strong case for the feasibility of peer-assisted rendering through the following two messages. First, by analyzing a large number of user traces from a popular virtual world called Second Life, we show that there are surprisingly many users with similar viewpoints and encompass large number of common objects in their viewing areas, indicating that a client can potentially find multiple others peers that can assist in rendering. Second, by combining three different rendering methods, each contributing to rendering of different classes of objects in the scene, we show that it is possible for a client to render the scene efficiently with little visual artifacts.

VORTEX participant: Géraldine Morin.

External participants: Wei Cheng, Wei Tsang Ooi, Minhui Zhu (National University of Singapore), Sébastien Mondet (University of Oslo, Norway).

Associated publication:

Minhui Zhu, Sébastien Mondet, Géraldine Morin, Wei Tsang Ooi, Wei Chen, *Towards peer-assisted rendering in networked virtual environments*, 19th ACM International Conference on Multimedia, pages 183-192, Scottsdale, USA, November/December 2011.

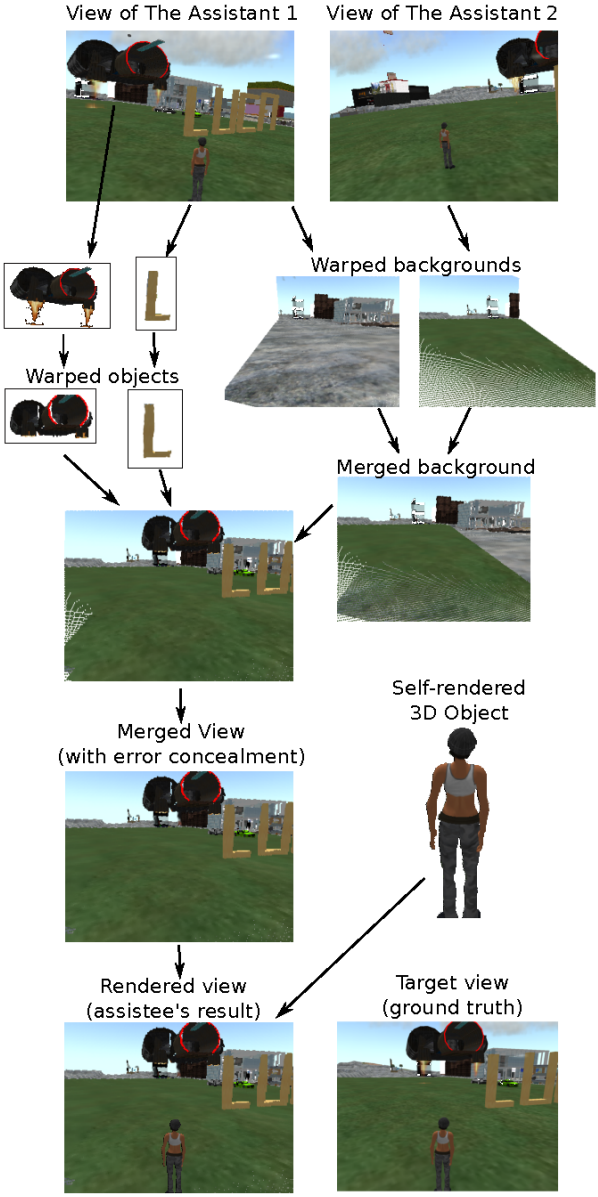


Illustration - Example of hybrid rendering.

ROM project

Keywords: Camera Tracking, Augmented Reality, Real-time Preview.

Real time On-set Matchmoving is an ANR (Agence Nationale de la Recherche) project in collaboration with the special effect company Duran-Duboi and the INRIA research center in Grenoble. This project aims at the creation of special effects and more specifically preview of these effects while shooting.

It is aimed at deleting boundary between production and post production by enabling the special effect director to preview the final compositing result during the shooting.

The main technological challenge is to provide to the classical camera tracking problem a solution that would control the trade-off between precision and computing time.

The original approach used in this project is based on three key elements:

- Variety of shooting scenes (indoor, outdoor, studio, etc.): preliminary phase of preparation to the end of which the stage can be equipped and enriched with artificial targets in order to ensure robustness and accuracy for tracking.

- Preparation phase: creation of a database of 2D and 3D features from which a subset could be tracked correctly and in real time.

- Real-time preview: using of this database and an innovative architecture combining software and recent graphics processing units (GPU).

VORTEX participants: Lilian Calvet, Vincent Charvillat, Pierre Gurdjos, Géraldine Morin.

Partners: Duran Duboi company, Perception team (INRIA Grenoble Rhône-Alpes).

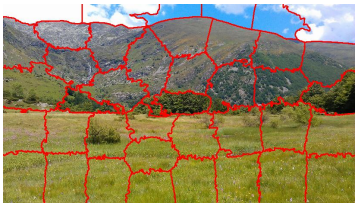
SICASSE project

Keywords: Segmentation, Superpixel, Image Registration.

The main SICASSE project objective is to provide to environmentalists an interactive tool for monitoring the evolution of Regional Natural parks over seasons and years. An image database is created by scientists and hikers: via a mobile application, they are led by GPS to "ecological interest points", and are then asked to take a picture of the landscape. Our first contribution in this project is to ensure that the picture taken is correct, by adding an augmented reality layer on the smartphone camera to guide the user while taking the photo. If the picture is correct, it is sent to the main server application to be processed. Here is our second contribution: the received image is then registred in comparison to previous photos of the same landscape. The aligned image is then segmented via a divide and merge method, using superpixels, and this data is then forwarded to a statistical analysis module which will monitor the evolution of the environment.

VORTEX participants: Benjamin Ahsan, Vincent Charvillat, Jérôme Guénard, Zouheir Hamrouni.

Partners: OKTAL-SE company, M3 Systems company, SEEM (Station d'Écologie Expérimentale du CNRS à Moulis), PNRPA (Parc Naturel Régional des Pyrénées Ariégeoises).



Illustrations - Left: guiding the user while taking picture. Right: an example of a superpixel grid.

SODIMEL project

Keywords: 3D Reconstruction, Quadrics of Revolution.

This project is in collaboration with the SODIMEL company whose goal is to produce wine-producing equipment. SODIMEL is trying to create a device able to predict the quantity of wine produced by a given area. It has been shown by the "Institut Francais de la Vigne et du Vin" that there exists a correlation between this output and the average volume of the grape two months before the vintage. A device has been developped to compute the volume of a grape without needing to pick it up. However, the measuring errors were too large to consider commercializing this tool. We have improved the system as to reduce the measuring errors, by reconstructing the grape in 3D. Our proposed device is made of two cameras with an angle of 30 degrees and a support holding the grape. The goal is then to estimate the volume of the grape from its images. We first calibrate the two views with patterns composed of concentric circles. The grape is assumed to be a set of prolate ellipsoids of revolution. We reconstruct the visible part of the grape. Eventually, using statistical models we estimate the volume, with a very low estimation error.

VORTEX participants: Vincent Charvillat, Jérôme Guénard, Pierre Gurdjos.

Partners: IFV (Institut Français de la Vigne et du Vin), SODIMEL company.

Example - Our proposed device and an example of grape reconstruction.

VINNEO project

Keywords: 3D Reconstruction, Tracking, Modeling, Calibration.

VINNEO is a project which brings together actors for wine production in the South-West of France. Our contribution is to compute some useful characteristics of the grapeyard, e.g. the leaves density or the average size of a vine on a parcel. Our goal is to infer these characteristics from a visual analysis. In order to do that, we have acquired videos passing through the rows of a vineyard. A first step has been to reconstruct a set of 3D points. Using these data, and assumptions about plant models and vineyards, we would like to reconstruct a 3D model of the visible parts. We are then hoping to be able to deduce interesting information.

VORTEX participants: Vincent Charvillat, Jérôme Guénard, Pierre Gurdjos, Géraldine Morin.

Partner: IFV (Institut Français de la Vigne et du Vin).

Example - On the left, a rectified image of a vine. On the right, a 3D reconstruction of this vine.

Publications 2011

International journals:

Wei Cheng, Wei Tsang Ooi, Sébastien Mondet, Romulus Grigoras, Géraldine Morin, *Modeling Progressive Mesh Streaming: Does Data Dependency Matter?*, ACM Transactions on Multimedia Computing, Communications, and Applications 7(2), February 2011.

International conferences and workshops:

Roberto Mecca, Jean-Denis Durou, *Unambiguous Photometric Stereo Using Two Images*, 16th International Conference on Image Analysis and Processing (ICIAP 2011), Ravenna, Italy, September 2011.

Axel Carlier, Arash Shafiei, Julien Badie, Salim Bensiali, Wei Tsang Ooi, *COZI: Crowdsourced and Content-based Zoomable Video Player*, 19th ACM International Conference on Multimedia, pages 829-830, Scottsdale, USA, November/December 2011.

Axel Carlier, Ravindra Guntur, Vincent Charvillat, Wei Tsang Ooi, *Combining Content-based Analysis and Crowdsourcing to Improve User Interaction with Zoomable Video*, 19th ACM International Conference on Multimedia, pages 43-52, Scottsdale, USA, November/December 2011.

Minhui Zhu, Sébastien Mondet, Géraldine Morin, Wei Tsang Ooi, Wei Chen, *Towards peer-assisted rendering in networked virtual environments*, 19th ACM International Conference on Multimedia, pages 183-192, Scottsdale, USA, November/December 2011.

National conferences and workshops:

Axel Carlier, Vincent Charvillat, *À propos d'interactions qui permettent d'analyser une vidéo*, ORASIS 2011, Praz-sur-Arly, June 2011.

Lilian Calvet, Pierre Gurdjos, Vincent Charvillat, *Suivi de caméra à partir de marqueurs plans composés de cercles concentriques : paradigme et algorithmes*, ORASIS 2011, Praz-sur-Arly, June 2011.

Jérôme Guénard, Géraldine Morin, Frédéric Boudon, Pierre Gurdjos, Vincent Charvillat, *Reconstruction de modèles virtuels de vignes à partir d'images*, ORASIS 2011, Praz-sur-Arly, June 2011.

Viorica PĂTRĂUCEAN, Pierre GURDJOS, Géraldine MORIN, Jean CONTER, *Détection de primitives linéaires ou circulaires par une approche a contrario*, ORASIS 2011, Praz-sur-Arly, June 2011.

Jean-Denis DUROU, Vincent CHARVILLAT, Pierre GURDJOS, Maxime DARAMY, *Résolution du shape-from-shading par apprentissage*, ORASIS 2011, Praz-sur-Arly, June 2011.

Jean-Denis DUROU, Xavier DESCOMBES, Pavel LUKASHEVISH, Aliaxandr KRAUSHONAK, *Reconstruction 3D du bâti à partir d'une seule image par naissances et morts multiples*, GRETSI 2011, Bordeaux, September 2011.

Summary

IRIT (UMR CNRS 5505)
VORTEX Team
(Visual Objects from Reality To EXpression)

ENSEEIHT
2, rue Camichel, BP 7122
31071 TOULOUSE Cedex 7
05.34.32.21.71

Front page illustration:
Real-time tracking using a point-based 3D model (toy leopard) acquired by a range scanner.
VORTEX contributors : C. Dehais, G. Morin and V.Charvillat.
(dehais.perso.enseeiht.fr/pbmtracking)