Emerging Trends in Visual Computing (ETVC'08)

LIX Colloquium 2008
École Polytechnique
Palaiseau, France

18th-20th November 2008 (last updated October 13, 2008)

Shun-ichi AMARI
Mathematical Neuroscience Laboratory, Brain Science Institute, RIKEN, Wako-Shi, Japan

Information Geometry and Its Applications:
Information geometry emerged from studies on invariant properties of a manifold of probability distributions. It includes convex analysis and its duality as a special but important part. Here, we begin with a convex function, and construct a dually flat manifold. The manifold possesses a Riemannian metric, two types of geodesics, and a divergence function. The generalized Pythagorean theorem and dual projections theorem are derived therefrom.We construct alpha-geometry, extending this convex analysis. In this review, geometry of a manifold of probability distributions is then given, and a plenty of applications are touched upon. Appendix presents an easily understable introduction to differential geometry and its duality.

Tetsuo ASANO
School of Information Science, Japan Advanced Institute of Science and Technology, JAIST, Japan

Constant-Working-Space Algorithms for Image Processing:
This talk surveys recent progress in constant-working-space algorithms for problems related to image processing. An extreme case is when an input image is given as read-only memory in which reading an array element is allowed but writing any value at any array element is prohibited, and also the number of working storage cells available for algorithms is at most some constant. This chapter shows how a number of important fundamental problems can be solved in such a highly constrained situation.

Francis BACH
INRIA/ENS, France

Machine learning and kernel methods for computer vision:
Kernel methods are a new theoretical and algorithmic framework for machine learning. By representing data through well defined dot-products, referred to as kernels, they allow to use classical linear supervised machine learning algorithms to non linear settings and to non vectorial data. A major issue when applying these methods to image processing or computer vision is the choice of the kernel. I will present recent advances in the design of kernels for images that take into account the natural structure of images.

Frédéric BARBARESCO
Thales Air Systems, France

Applications of Information Geometry to Radar Signal Processing:
Main issue of High Resolution Doppler Imagery is related to robust statistical estimation of Toeplitz Hermitian positive definite covariance matrices of sensor data time series (e.g. in Doppler Echography, in Underwater acoustic, in Electromagnetic Radar, in Pulsed Lidar). We consider this problem jointly in the framework of Riemannian symmetric spaces and the framework of Information Geometry. Both approaches lead to the same metric, that has been initially considered in other mathematical domains (study of Bruhat-Tits complete metric Space & Upper-half Siegel Space in Symplectic Geometry). Based on Fréchet-Karcher barycenter definition & geodesics in Bruhat-Tits space, we address problem of N Covariance matrices Mean estimation. Our main contribution lies in the development of this theory for Complex Autoregressive models (maximum entropy solution of Doppler Spectral Analysis). Specific Blocks structure of the Toeplitz Hermitian covariance matrix is used to define an iterative & parallel algorithm for Siegel metric computation. Based on Affine Information Geometry theory, we introduce for Complex Autoregressive Model, Kähler metric on reflection coefficients based on Kähler potential function given by Doppler signal Entropy. The metric is closely related to Kähler-Einstein manifold and complex Monge-Ampere Equation. Finally, we study geodesics in space of Khler potentials and action of Calabi & Kähler-Ricci Geometric Flows for this Complex Autoregressive Metric. We conclude with different results obtained on real Doppler Radar Data in HF & X bands : X-band radar monitoring of wake vortex turbulences, detection for Coastal X-band & HF Surface Wave Radars.

Michel BARLAUD
I3S CNRS, University of Nice-Sophia-Antipolis, Polytech'Nice & Institut Universitaire de France, France

Image Retrieval via Kullback Divergence of Patches of Wavelets Coefficients in the k-NN Framework:
This talk presents a framework to define an objective measure of the similarity (or dissimilarity) between two images for image processing. The problem is twofold: 1) define a set of features that capture the information contained in the image relevant for the given task and 2) define a similarity measure in this feature space. In this paper, we propose a feature space as well as a statistical measure on this space. Our feature space is based on a global description of the image in a multiscale transformed domain. After decomposition into a Laplacian pyramid, the coefficients are arranged in intrascale/ interscale/interchannel patches which reflect the dependencies of neighboring coefficients in presence of specific structures or textures. At each scale, the probability density function (pdf) of these patches is used as a description of the relevant information. Because of the sparsity of the multiscale transform, the most significant patches, called Sparse Multiscale Patches (SMP), describe efficiently these pdfs. We propose a statistical measure (the Kullback-Leibler divergence) based on the comparison of these probability density function. Interestingly, this measure is estimated via the nonparametric, k-th nearest neighbor framework without explicitly building the pdfs. This framework is applied to a query-by-example image retrieval method. Experiments on two publicly available databases showed the potential of our SMP approach for this task. In particular, it performed comparably to a SIFT-based retrieval method and two versions of a fuzzy segmentation-based method (the UFM and CLUE methods), and it exhibited some robustness to different geometric and radiometric deformations of the images.

Jean-Daniel BOISSONNAT
GEOMETRICA, INRIA Sophia-Antipolis, France

Certified Mesh Generation:
Given a domain D, the problem of mesh generation is to construct a simplicial complex that approximates D in both a topological and a geometrical sense and whose elements satisfy various constraints such as size, aspect ratio or anisotropy. The talk will cover some recent results on triangulating surfaces and volumes by Delaunay refinement, anisotropic mesh generation and surface reconstruction. Applications in medical images, computer vision and geology will be discussed.

Pascal FUA
EPFL, CVLAB, Swiss

Recovering Shape and Motion from Video Sequences:
In recent years, because cameras have become inexpensive and ever more prevalent, there has been increasing interest in video-based modeling of shape and motion. This has many potential applications in areas such as electronic publishing, entertainment, sports medicine and athletic training. It, however, is an inherently difficult task because the image-data is often incomplete, noisy, and ambiguous. In our work, we focus on the recovery of deformable and articulated 3D motion from single video sequences. In this talk, I will present the models we have developed for this purpose and demonstrate the applicability of our technology for Augmented Reality and human body tracking purposes. Finally, I will present some open research issues and discuss our plans for future developments.

Markus GROSS
Department of Computer Science, Institute of Scientific Computing, Swiss Federal Institute of Technology Zurich, ETHZ, Switzerland

3D Video: A Fusion of Graphics and Vision:
In recent years 3-dimensional video has received a significant attention both in research and in industry. Applications range from special effects in feature films to the analysis of sports events. 3D video is concerned with the computation of virtual camera positions and fly-throughs of a scene given multiple, conventional 2D video streams. The high-quality synthesis of such view-independent video representations poses a variety of technical challenges including acquisition, reconstruction, processing, compression, and rendering. In this talk I will outline the research in this area carried out at ETH over the past years. I will discuss various concepts for passive and active acquisition of 3D video using combinations of multiple cameras and projectors. Furthermore, I will address topics related to the representation and processing of the massive amount data arising from such multiple video streams. I will highlight the underlying technical concepts and algorithms that draw upon knowledge both from graphics and from vision. Finally I will demonstrate some commercial applications targeting at virtual replays for sports broadcasts.

Xianfeng David GU
State University of New York at Stony Brook, USA

Discrete Curvature Flow for Surfaces and 3-Manifolds:
This talk introduce the concepts, theories and algorithms for discrete curvature flows for surfaces with arbitrary topologies. Discrete curvature flow for hyperbolic 3-manifolds with geodesic boundaries are also explained. Curvature flow method can be used to design Riemannian metrics by prescribed curvatures, and applied for parameterization in graphics, shape registration and comparison in vision and brain mapping in medical imaging, spline construction in computer aided geometric design, and many other engineering fields.

Leonidas GUIBAS
Computer Science Department, Stanford University, USA

Detection of Symmetries and Repeated Patterns in 3D Point Cloud Data:
Digital models of physical shapes are becoming ubiquitous in our economy and life. Such models are sometimes designed ab initio using CAD tools, but more and more often they are based on existing real objects whose shape is acquired using various 3D scanning technologies. In most instances, the original scanner data is just a set, but a very large set, of points sampled from the surface of the object. We are interested in tools for understanding the local and global structure of such large-scale scanned geometry for a variety of tasks, including model completion, reverse engineering, shape comparison and retrieval, shape editing, inclusion in virtual worlds and simulations, etc. This talk will present a number of point-based techniques for discovering global structure in 3D data sets, including partial and approximate symmetries, shared parts, repeated patterns, etc. It is also of interest to perform such structure discovery across multiple data sets distributed in a network, without actually ever bring them all to the same host.

Sylvain LAZARD
VEGAS, INRIA LORIA Nancy, France

3D Visibility and Lines in Space:
Computing visibility information in a 3D environment is crucial to many applications such as computer graphics, vision and robotics. Typical visibility problems include computing the view from a given point, determining whether two objects partially see each other, and computing the umbra and penumbra cast by a light source. In a given scene, two points are visible if the segment joining them does not properly intersect any obstacle in the scene. The study of visibility is thus intimately related to the study of the set of free line segments in a scene. In this talk, I will review some recent combinatorial and algorithmic results related to non-occluded segments tangent to up to four objects in three dimensional scenes.

Stéphane MALLAT
École Polytechnique, Centre de Mathématiques Appliquées (CMAP), France

Sparse Geometric Super-Resolution:
What is the maximum signal resolution that can be recovered from partial noisy or degraded data? This inverse problem is a central issue, from medical to satellite imaging, from geophysical seismic to HDTV visualization of Internet videos. Increasing an image resolution is possible by taking advantage of "geometric regularities", whatever it means. Super-resolution can indeed be achieved for signals having a sparse representation which is "incoherent" relatively to the measurement system. For images and videos, it requires to construct sparse representations in redundant dictionaries of waveforms, which are adapted to geometric image structures. Signal recovery in redundant dictionaries is discussed, and applications are shown in dictionaries of bandlets for image super-resolution.

Hiroshi MATSUZOE
Department of Computer Science and Engineering Graduate School of Engineering, Nagoya Institute of Technology, NITECH, Japan

Computational Geometry from the Viewpoint of Affine Differential Geometry:
Incidence relations (configurations of vertexes, edges, etc.) are important in computational geometry. Incidence relations are invariant under the group of affine transformations. On the other hand, affine differential geometry is to study hypersurfaces in an affine space that are invariant under the group of affine transformation. Therefore affine differential geometry gives a new sight in computational geometry. From the viewpoint of affine differential geometry, algorithms of geometric transformation and dual transformation are discussed. The Euclidean distance function is generalized by a divergence function in affine differential geometry. A divergence function is an asymmetric distance-like function on a manifold, and it is an important object in information geometry. For divergence functions, the upper envelope type theorems on statistical manifolds are given. Voronoi diagrams determined from divergence functions are also discussed.

Dimitris METAXAS
Computational Biomedicine Imaging and Modeling Center, CBMI, Rutgers University, USA

Unifying Subspace and Distance Metric Learning with Bhattacharyya Coefficient for Image Classification:
In this talk, we propose a unified scheme of subspace and distance metric learning under the Bayesian framework for image classification. According to the local distribution of data, we divide the k-nearest neighbors of each sample into the intra-class set and the inter-class set, and we aim to learn a distance metric in the embedding subspace, which can make the distances between the sample and its intra-class set smaller than the distances between it and its inter-class set. To reach this goal, we consider the intra-class distances and the inter-class distances to be from two different probability distributions respectively, and we model the goal with minimizing the overlap between two distributions. Inspired by the Bayesian classification error estimation, we formulate the objective function by minimizing the Bhattachyrra coefficient between two distributions. We further extend it with the kernel trick to learn nonlinear distance metric. The power and generality of the proposed approach are demonstrated by a series of experiments on the CMU-PIE face database, the extended YALE face database, and the COREL-5000 nature image database.

Frank NIELSEN
LIX, École Polytechnique, Paris, France & Sony Computer Science Laboratories Inc., Tokyo, Japan

Computational Geometry in Dually Flat Spaces: Theory, Applications and Perspectives:
Computational information geometry emerged from the fruitful interactions of geometric computing with information geometry. In this talk, we survey the recent results obtained in that direction by first describing generalizations of core algorithms of computational geometry and machine learning to broad and versatile classes of distortion measures. Namely, we introduce the generic classes of Bregman, Csiszar and Burbea-Rao parametric divergences and explain their relationships and properties with respect to algorithmic design. We then present few applications of these meta-algorithms to the field of statistics and data analysis and conclude with perspectives.

Richard NOCK
CEREGMIA, University of Antilles-Guyane, France

The Intrinsic Geometries of Learning:
In a seminal paper, Amari (1998) proved that learning can be made more efficient when one uses the intrinsic Riemanian structure of the algorithms' spaces of parameters to point the gradient towards better solutions. In this paper, we show that many learning algorithms, including various boosting algorithms for linear separators, the most popular top-down decision-tree induction algorithms, and some on-line learning algorithms, are spawns of a generalization of Amari's natural gradient to some particular non-Riemanian spaces. These algorithms exploit an intrinsic dual geometric structure of the space of parameters in relationship with particular integral losses that are to be minimized. We unite some of them, such as AdaBoost, additive regression with the square loss, the logistic loss, the top-down induction performed in CART and C4.5, as a single algorithm on which we show general convergence to the optimum and explicit convergence rates under very weak assumptions. As a consequence, many of the classification calibrated surrogates of Bartlett et al. (2006) admit efficient minimization algorithms.

Nikos PARAGIOS
École Centrale de Paris, ECP, Paris, France

Procedural Modeling of Architectures: Towards Large Scale Visual Reconstruction:
Three-dimensional content is a novel modality used in numerous domains like navigation, post production & cinematography, architectural modeling and urban planning. These domains have benefited from the enormous progress has been made on 3D reconstruction from images. Such a problem consists of building geometric models of the observed environment. State of the art methods can deliver excellent results in a small scale but suffer from being local and cannot be considered in a large scale reconstruction process since the assumption of recovering images from multiple views for an important number of buildings is rather unrealistic. On the other hand several efforts have been made in the graphics community towards content creation with city engines. Such models are purely graphics-based and given a set of rules (grammars) as well as dictionary of architectures (buildings) can produce virtual cities. Such engines could become far more realistic through the use of actual city models as well as knowledge of building architectures. Developing 3D models/rules/grammars that are image-based and coupling these models with actual observations is the greatest challenge of urban modeling. Solving the large-scale geometric modeling problem from minimal content could create novel means of world representation as well as novel markets and applications. In this talk, we will present some preliminary results on large scale modeling and reconstruction through architectural grammars.

Xavier PENNEC
ASCLEPIOS, INRIA Sophia-Antipolis, France

Statistical Computing on Manifolds for Computational Anatomy:
Computational anatomy is an emerging discipline that aims at analyzing and modeling the individual anatomy of organs and their biological variability across a population. The goal is not only to model the normal variations among a population, but also discover morphological differences between normal and pathological populations, and possibly to detect, model and classify the pathologies from structural abnormalities. Applications are very important both in neuroscience, to minimize the influence of the anatomical variability in functional group analysis, and in medical imaging, to better drive the adaptation of generic models of the anatomy (atlas) into patient-specific data (personalization). However, understanding and modeling the shape of organs is made difficult by the absence of physical models for comparing different subjects, the complexity of shapes, and the high number of degrees of freedom implied. Moreover, the geometric nature of the anatomical features usually extracted raises the need for statistics and computational methods on objects that do not belong to standard Euclidean spaces. We investigate in this chapter the Riemannian metric as a basis for developing generic algorithms to compute on manifolds. We show that few computational tools derived from this structure can be used in practice as the atoms to build more complex generic algorithms such as mean computation, Mahalanobis distance, interpolation, filtering and anisotropic diffusion on fields of geometric features. This computational framework is illustrated with the joint estimation and anisotropic smoothing of diffusion tensor images and with the modeling of the brain variability from sulcal lines.

Ramesh RASKAR
MIT Media Lab, USA

Computational Photography: Epsilon to Coded Imaging:
Computational photography combines plentiful computing, digital sensors, modern optics, actuators, and smart lights to escape the limitations of traditional cameras, enables novel imaging applications and simplifies many computer vision tasks. However, a majority of current Computational Photography methods involve taking multiple sequential photos by changing scene parameters and fusing the photos to create a richer representation. The goal of Coded Computational Photography is to modify the optics, illumination or sensors at the time of capture so that the scene properties are encoded in a single (or a few) photographs. We describe several applications of coding exposure, aperture, illumination and sensing and describe emerging techniques to recover scene parameters from coded photographs.

Cordelia SCHMID
LEAR, INRIA Grenoble, France

Large-Scale Object Recognition Systems:

This paper introduces recent methods for large scale image search. State-of-the-art methods build on the bag-of-features image representation. We first analyze bag-of-features in the framework of approximate nearest neighbor search. This shows the sub-optimality of such a representation for matching descriptors and leads us to derive a more precise representation based on 1) Hamming embedding (HE) and 2) weak geometric consistency constraints (WGC). HE provides binary signatures that refine the matching based on visual words. WGC filters matching descriptors that are not consistent in terms of angle and scale. HE and WGC are integrated within the inverted file and are efficiently exploited for all images, even in the case of very large datasets. Experiments performed on a dataset of one million of images show a significant improvement due to the binary signature and the weak geometric consistency constraints, as well as their efficiency. Estimation of the full geometric transformation, i.e., a re-ranking step on a short list of images, is complementary to our weak geometric consistency constraints and allows to further improve the accuracy.

Gabriel TAUBIN
Division of Engineering, Brown University, USA

Shape from Depth Discontinuities:

We propose a new primal-dual framework for representation, capture, processing, and display of piecewise smooth surfaces, where the dual space is the space of oriented 3D lines, or rays, as opposed to the traditional dual space of planes. An image capture process detects points on a depth discontinuity sweep from a camera moving with respect to an object, or from a static camera and a moving object. A depth discontinuity sweep is a surface in dual space composed of the time-dependent family of depth discontinuity curves span as the camera pose describes a curved path in 3D space. Only part of this surface, which includes silhouettes, is visible and measurable from the camera. Locally convex points deep inside concavities can be estimated from the visible non-silhouette depth discontinuity points. Locally concave point laying at the bottom of concavities, which do not correspond to visible depth discontinuities, cannot be estimated, resulting in holes in the reconstructed surface. A first variational approach to fill the holes, based on fitting an implicit function to a reconstructed oriented point cloud, produces watertight models.We describe a first complete end-to-end system for acquiring models of shape and appearance.We use a single multi-flash camera and turntable for the data acquisition and represent the scanned objects as point clouds, with each point being described by a 3-D location, a surface normal, and a Phong appearance model.

Baba VEMURI
CISE Dept., University of Florida, USA

Information-Theoretic Algorithms for Diffusion Tensor Imaging:
Concepts from Information Theory have been used quite widely in Image Processing, Computer Vision and Medical Image Analysis for several decades now. Most widely used concepts are that of KL-divergence, minimum description length (MDL), etc. These concepts have been popularly employed for image registration, segmentation, classification etc. In this chapter we review several methods, mostly developed by our group at the Center for Vision, Graphics & Medical Imaging in the University of Florida, that glean concepts from Information Theory and apply them to achieve analysis of Diffusion-Weighted Magnetic Resonance (DW-MRI) data. This relatively new MRI modality allows one to non-invasively infer axonal connectivity patterns in the central nervous system. The focus of this chapter is to review automated image analysis techniques that allow us to automatically segment the region of interest in the DWMRI image wherein one might want to track the axonal pathways and also methods to reconstruct complex local tissue geometries containing axonal fiber crossings. Implementation results illustrating the algorithm application to real DW-MRI data sets are depicted to demonstrate the effectiveness of the methods reviewed.

Suresh VENKATASUBRAMANIAN
School of Computing, University of Utah, USA

Non-standard Geometries and Data Analysis:
Traditional data mining starts with the mapping from entities to points in a Euclidean space. The search for patterns and structure is then framed as a geometric search in this space. Concepts like principal component analysis, regression, clustering, and centrality estimation have natural geometric formulations, and we now understand a great deal about manipulating such (typically high dimensional) spaces. For many domains of interest however, the most natural space to embed data in is not Euclidean. Data might lie on curved manifolds, or even inhabit spaces endowed with different distance structures than l_p spaces. How does one do data analysis in such domains? In this talk, I'll discuss two specific domains of interest that pose challenges for traditional data mining and geometric methods. One space consists of collections of distributions, and the other is the space of shapes. In both cases, I'll present ongoing work that attempts to interpret and understand clustering in such spaces, driven by different applications.

Martin VETTERLI
School of Computer and Communication Sciences, EPFL, Switzerland

Sparse Sampling: Variations on a Theme by Shannon:
Sampling is not only a beautiful topic in harmonic analysis, with an interesting history, but also a subject with high practical impact, at the heart of signal processing and communications and their applications. The question is very simple: when is there a one-to-one relationship between a continuous-time function and adequately acquired samples of this function? A cornerstone result is of course Shannon's sampling theorem, which gives a sufficient condition for reconstructing the projection of a signal onto the subspace of bandlimited functions, and this by taking inner products with a sinc function and its shifts. Many variations of this basic framework exist, and they are all related to a subspace structure of the classes of objects that can be sampled. Recently, this framework has been extended to classes of non-bandlimited sparse signals, which do not have a subspace structure. Perfect reconstruction is possible based on a suitable projection measurement. This gives a sharp result on the sampling and reconstruction of sparse continuous-time signals, namely that 2K measurements are necessary and sufficient to perfectly reconstruct a K-sparse continuous-time signal. In accordance with the principle of parsimony, we call this sampling at Occam's rate. We first review this result and show that it relies on structured Vandermonde measurement matrices, of which the Fourier matrix is a particular case. It also uses a separation into location and value estimation, the first being non-linear, while the second is linear. Because of this structure, fast, $O(K^3)$ methods exist, and are related to classic algorithms used in spectral estimation and error correction coding. We then generalize these results to a number of cases where sparsity is present, including piecewise polynomial signals, as well as to broad classes of sampling or measurement kernels, including Gaussians and splines. Of course, real cases always involve noise, and thus, retrieval of sparse signals in noise is considered. That is, is there a stable recovery mechanism, and robust practical algorithms to achieve it. Lower bounds by Cramer-Rao are given, which can also be used to derive uncertainty relations with respect to position and value of sparse signal estimation. Then, a concrete estimation method is given using an iterative algorithm due to Cadzow, and is shown to perform close to optimal over a wide range of signal to noise ratios. This indicates the robustness of such methods, as well as their practicality. Next, we consider the connection to compressed sensing and compressive sampling, a recent approach involving random measurement matrices, a discrete set up, and retrieval based on convex optimization. These methods have the advantage of unstructured measurement matrices (actually, typically random ones) and therefore a certain universality, at the cost of some redundancy. We compare the two approaches, highlighting differences, similarities, and respective advantages. Finally, we move to applications of these results, which cover wideband communications, noise removal, and superresolution imaging, to name a few. We conclude by indicating that sampling is alive and well, with new perspectives and many interesting recent results and developments. Joint work with Thierry Blu (CUHK), Lionel Coulot, Ali Hormati (EPFL), Pier-Luigi Dragotti (ICL) and Pina Marziliano (NTU).

Jun ZHANG
Department of Psychology, University of Michigan, USA

Information Geometry: Duality, Convexity and Divergences:
In this talk, I explore the mathematical relationships between duality in information geometry, convex analysis, and divergence functions. First, from the fundamental inequality of a convex function, a family of divergence measures can be constructed, which specializes to the familiar Bregman divergence, Jenson difference, beta-divergence, and alpha-divergence, etc. Second, the mixture parameter turns out to correspond to the alpha $\leftrightarrow$ -alpha duality in information geometry (which I call "referential duality", since it is related to the choice of a reference point for computing divergence). Third, convex conjugate operation induces another kind of duality in information geometry, namely, that of biorthogonal coordinates and their transformation (which I call "representational duality", since it is related to the expression of geometric quantities, such as metric, affine connection, curvature, etc of the underlying manifold). Under this analysis, what is traditionally called "+1/-1 duality" and "e/m duality" in information geometry reflect two very different meanings of duality that are nevertheless intimately interwined for dually flat spaces.