Optimal Mass Transport for Understanding and Synthesis of Visual Data
PIYRA proposal 2007 by Dr. Rozenn DAHYOT

Proposed Research Objectives

The optimal mass transport problem, introduced by G. Monge in 1781, poses the question of moving a pile of soil from one site to another while minimizing the transportation cost. Since, it has been reformulated by Kantorovitch in 1948 (and hence renamed the Monge-Kantorovitch problem), and it is used in many areas of sciences such as econometrics or civil engineering (design of transportation or water network). As the domain of applications of this method is very broad, it is currently a problem of major interest in Mathematics and several numerical solutions have recently been proposed. This has also been applied in particular to medical image registration which puts in correspondence images (of brains for instance) of one patient at different times or of different patients. This allows to align all images to enhance and detect abnormalities.

As one objective, this proposal aims to apply the optimal mass transport formalism in new computer vision applications.

- The first application considers the computation of the transfer function (a.k.a warping or registration function) between two successive images in a video. In particular, we want to explore the applications of the transfer function as a representation of an object deformation (e.g. face) to analyse its dynamic state (e.g. facial emotion). This can be used for instance, for video analysis and indexing, human-computer interactions or video surveillance. Another application sought proposes to apply this modelling to synthesis for computer graphics (e.g. face animation in video games, or for film special effects).
- The second will investigate the computation of the transfer function to match two continuous multidimensional probability density functions in the context of offline visual tracking. To track simultaneously and robustly several objects (or several part of an elastic object such as a human body) in a video, is a challenge as many difficulties such as occlusions or missing data can occur. This proposal is in continuation with our work on multiple object tracking using the Viterbi algorithm, where the formalism is discrete. We will then extend this scheme from discrete probability density functions to continuous ones, and consequently improve robustness of the whole process.

Advances in the computer vision field often happen when complex theoretical framework from Statistics and Mathematics are made easily understandable and available to the researchers. The recent developments on the Monge-Kantorovitch problem in Mathematics show that high value research outputs in applied science will come in the next few years.

In parallel, in Statistics, copulas have recently become a huge subject of interest. Copulas are *functions that joins or couple multivariate distribution functions to their one-dimensional marginal distributions*. In particular, for the two dimensional case, a copula expresses the joint density function with respect to its two one-dimensional marginals. Recent numerical solutions to the optimal mass transport problem match multivariate distributions using their one-dimensional marginals. In particular, we have proposed recently to compute the transfer function using kernel based representations of the marginals. This has been successfully applied to transfer the colour statistics from one picture to another (application to film editing). However the link in between the two frameworks (Monge-Kantorovitch problem and copulas) is yet to be shown. Another major objective of this research consists in improving the clarity of the algorithms considering the recent developments in statistics on Copulas, in image processing and Mathematics.

To conclude, this proposal aims at:

- deepening and clarifying the understanding and the application of the Monge-Kantorovitch formalism.
- applying this formalism to new applications,
- proposing algorithms that give accurate results for a minimum computation time.

This research will therefore benefit to a wide community of researchers. For illustration, we will consider two main applications in computer vision:

- multiple object tracking in videos,
- object deformation analysis (e.g. face).

Those are important in applications such as video surveillance (e.g. airports, road traffics, etc.), parsing (e.g. for filtering out illicit video for internet), human machine interfaces, and computer graphics (e.g. game, special effect, etc.).