Photographie computationnelle : copier/coller d'objets vidéos (Proposition de stage X2005 — 2007)

Frank NIELSEN
(E-mail :Frank.Nielsen@acm.org)

Thématique	Traitement d'image, algorithmique, interface homme/machine
Laboratoire	LIX, École Polytechnique, Paris
Durée	4 à 5 mois (Avril 2008 ∼, anglais ou français)

Positionnement du stage

frame 9 frame 26 frame 39 frame 53

L'avénement des appareils photos et caméscopes numériques a permis outre le fait de copier parfaitement bit par bit le support (en s'affranchissant de l'étape de digitalisation) la possibilité d'édition du contenu à la fois par les experts et les néophytes (non-linear editing systems). Bien que des outils puissants existent dans le cadre de la manipulation d'images (eg., GIMP, Adobe Photoshop®), il est encore difficile d'avoir un outil aussi convivial, rapide, et permettant le même degré de créativité pour les supports vidéos. Notamment, un des problèmes fondamentaux rencontré est la sélection dans une séquence vidéo d'un ou plusieurs objets animés dans une scène dynamique. Le but de ce stage est d'étudier quelques techniques récentes [1,2,3,4] permettant la sélection d'un objet vidéo et la copie de celui-ci dans une autre source vidéo (paste). Ce stage permettra de se familiariser avec un ensemble de techniques modernes en photographie computationnelle (traitement avancé de l'image : image gradient, poisson solver, tracking), et permettra de construire une brique de base ergonomique en s'intéressant aux problèmes d'interface homme/machine pour un logiciel de type videoshop [3].

Objectifs du stage

Le but de stage est de partir de la bibliographie [1,2,3,4] pour implanter un système permettant de copier et coller dans une autre vidéo un objet animé d'une vidéo source de manière la plus intuitive possible, en minimisant le nombre d'intéractions à la souris de l'utilisateur. Traditionnellement, les logiciels cherchent à copier le plus précisement possible un objet dans une image en créant un masque de transparence associé à l'objet (alpha, RGBA) : cela pose plusieurs problèmes de définir un bord soft et de tenir compte des partie semi-transparentes de l'objet (par exemple, des fenêtres). Le problème est en général défini de manière ambigüe, et plusieurs méthodes de segmentation efficaces existent dans le cadre d'un objet (cf. méthodes statistiques Bayésiennes [4], etc.)

sources/destinations cloning seamless cloning

S'il s'agit de coller un objet d'une image source vers une image destination, il existe une autre approche radicale et étonnament efficace qui consiste à entourer grossièrement l'objet à copier puis à le faire "fondre" dans l'image de destination. Ce processus est appelé : édition de type Poisson, puisqu'il travaille dans le gradient des images et cherche à reconstruire à partir d'une image gradient mixte l'image composite finale en résolvant un système d'équations par un solver de type Poisson [1,2,3]: $\Delta f = \mathrm{div}\mathbf{v}$ étant donné des conditions initiales (frontières), avec $\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$ (Laplacien) et $\mathrm{div}\mathbf{v} = \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y}$ (l'opérateur de divergence). Le stage comporte une forte composante de programmation en C++ ou Java afin de bâtir le système complet : (1) interface homme/machine, (2) sélection de primitive vidéos, et (3) collage dans une autre vidéo. Afin de ne pas recréer la roue, on se basera sur certains codes sources disponibles (eg., [2]).

Le stage permettra de regarder et de proposer du point de vue de la recherche :

- la gestion de la cohérence spatiale et temporelle,
- l'étude et la géneralisation du traitement d'image dans le domaine gradient. Par exemple, on regardera les pyramides de reconstruction successives à partir de différentielles successives d'images (gradient de gradient, etc.)
- le rôle de l'utilisateur dans la sélection, tracking, et correction d'imperfections dans le collage de vidéo clip-arts.

Profil/prérequis

Mots clefs: Solver Poisson, gradient image, reconstruction.
Outils: Java ou C++ (au choix).

Bibliographie

1. Yin Li, Jian Sun, Heung-Yeung Shum: Video object cut and paste. ACM Trans. Graph. 24(3): 595-600 (2005)

2. Patrick Perez, Michel Gangnet, Andrew Blake: Poisson image editing. ACM Trans. Graph. 22(3): 313-318 (2003). Code source en C++: `http://www.cs.tau.ac.il/~tommer/adv-graphics/ex1.htm`

3. Hongcheng Wang, Ning Xu, Ramesh Raskar, Narendra Ahuja: Videoshop: A new framework for spatio-temporal video editing in gradient domain. Graphical Models 69(1): 57-70 (2007), `http://vision.ai.uiuc.edu/~wanghc/research/editing.html`

4. Yung-Yu Chuang, Brian Curless, David Salesin, Richard Szeliski: A Bayesian Approach to Digital Matting. CVPR (2) 2001: 264-271

5. Frank Nielsen, *Visual Computing: Geometry, Graphics, and Vision*, ISBN 1584504277, Charles River Media, August 2005.