AVG
Project on “Structure-From-Motion”

1 Objectives

The main objective of thisproject is to understand the characteristics of a 3D reconstruction algorithm ("Structure-From-Motion") by monocular vision and to implement its main steps with Matlab or Python. The principle diagram is shown in Fig. 1.
A "Structure-From-Motion" algorithm can be broken down into 8 main steps :

1. Feature point detection : extract from the image a corner map (Harris detector) by calculating in each pixel the quantity :

$$R\left(x,y\right)=\det\left(\mathrm{M}\right)-k.\left[\mathrm{trace}\left(\mathrm{M}\right)\right]^{2} \tag{1}$$

with $k\in[0.04;0.06]$. M is the autocorrelation matrix of the image :

$$\mathrm{M}=\left[\begin{array}{cc}\sum\limits_{x,y\in W}w(x,y).\mathrm{I}_{x}^{2} & \sum\limits_{x,y\in W}w(x,y).\mathrm{I}_{x}.\mathrm{I}_{y}\\ \sum\limits_{x,y\in W}w(x,y).\mathrm{I}_{x}.\mathrm{I}_{y} & \sum\limits_{x,y\in W}w(x,y).\mathrm{I}_{y}^{2}\end{array}\right]=\left[\begin{array}{cc}A & C\\ C & B\end{array}\right] \tag{2}$$

where $w(x,y)=\frac{1}{2\pi\sigma^{2}}.e^{\frac{(x-x_{c})^{2}+(y-y_{c})^{2}}{2\sigma^{2}}}$ is a Gaussian mask centered on the W analysis window. Directional derivatives are approximated according to (optionally) Gradient, Prewitt or Sobel masks. You will be able to compare the different masks as well as different analysis window sizes.
2. Matching by correlation
   Correlation methods have long been used, particularly in photogrammetry, to map pixels based on intensity information. The idea is to define a measure of similarity between the pixels of two images, the principle is to consider, for a pixel p_1 of image 1, a square window centered on p_1 and to calculate its correlation/distance with a window in the second image.The correlation function is then maximum in p_2 corresponding to p_1 in the second image (minimum distance).
   The most common distance measurements are Sum of Absolute Differences (SAD), Sum of Squared Differences (SSD) or Zero-mean Normalized Cross-Correlation (ZNCC). It is also possible to use the feature descriptors to perform the matching using FLANN approches.
3. Robust estimation of the fundamental matrix or the essential matrix
4. Reconstruction of the projection matrices
5. Computation of the 3D points by triangulation
6. Bundle Adjustment
7. 3D shape generation

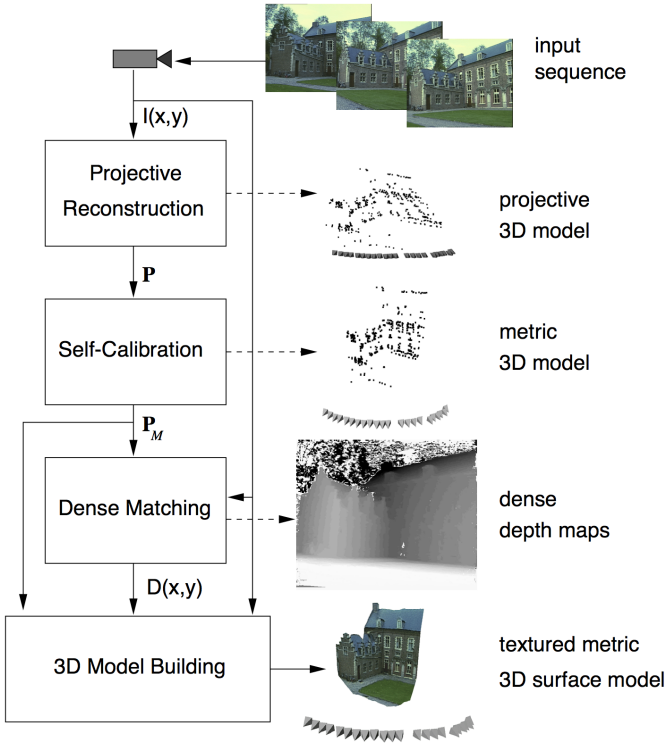

FIGURE 1 – Schéma de principe d'un algorithme de reconstruction 3D

2 To do

First of all, you need to find a video sequence. For the project, you will use the KITTI dataset (http ://www.cvlibs.net/datasets/kitti/raw_data.php).
You will then implement the above steps in Matlab or Python to obtain a 3D model of the observed scene. You are free to search the Internet, and you can use existing functions, but the evaluation criteria favour personal functions.

3 Deliverables

For the project, you will deliver an electronic version of a small report (2-3 pages) in pdf format by email, as well as the Matlab/Python code with comments. The evaluation criteria are based on :

— The level of success of the objectives (over the 8 steps)
— Customization of the code
— Calculation time