MATHÉMATIQUES
VISION
APPRENTISSAGE

MASTER MVA

VISION 3D ARTIFICIELLE

Report

GABRIEL ATHÈNES

November 10, 2021

Contents

1 Introduction

In this report I analyse the results of a 3D reconstruction algorithm based on the Graph Cut method. In order to use this method, we first create an oriented graph with nodes representing triplets (x,y,d) with (x,y) the coordinates of the pixels of the picture and d an associated disparity. The graph can therefore be represented like a cube in $\mathbb{N}^3$ with square layers corresponding to images shifted horizontally by a disparity d. A node (x,y,d) is connected to its neighboors in $\mathbb{N}^3$. The weights between neighbors (x,y,d) and $(x,y,d+1)$ correspond to the zero-normalized cross correlation between a patch around the pixels (x,y) of image 1 and a patch around the pixels $(x+d,y)$ of image 2. The weights between neighbors (x,y,d) and $(x+i,y,d)$ (resp. $(x,y+j,d)$) are fixed weights λ. Therefore the value of a cut of the graph depends of the correlation between every pair of patch around (x,y) and (x,y,d), as well as a positive term proportional to the difference of disparity between two neighbors. Indeed if (x,y) has disparity d and $(x+1,y)$ has disparity d', then a cut has to cut through $d'-d$ arcs between (x,y,d) and (x,y,d') which has a cost of $\lambda \times (d'-d)$. A cut therefore consists of a pixel to pixel correspondance term as well as a "neighborhood coherence" term. A minimum cut therefore realizes the best tradeoff between these two terms.

2 Results with the given images

Figure 1: Disparity map

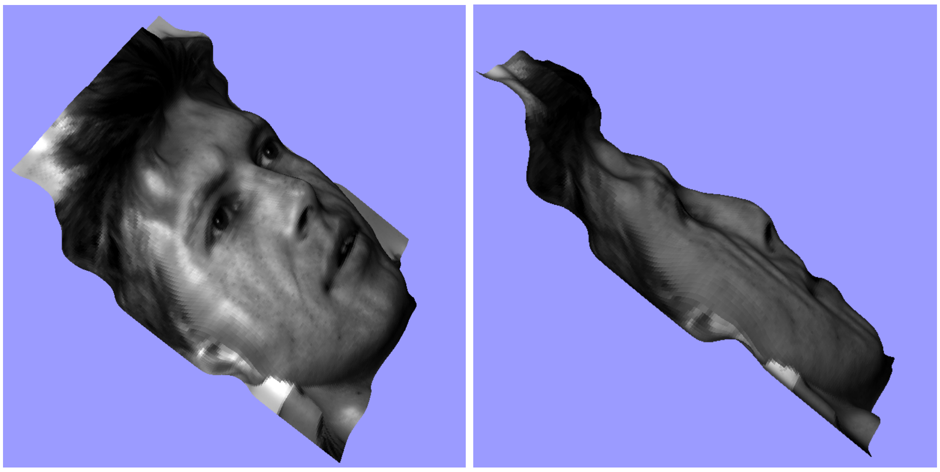
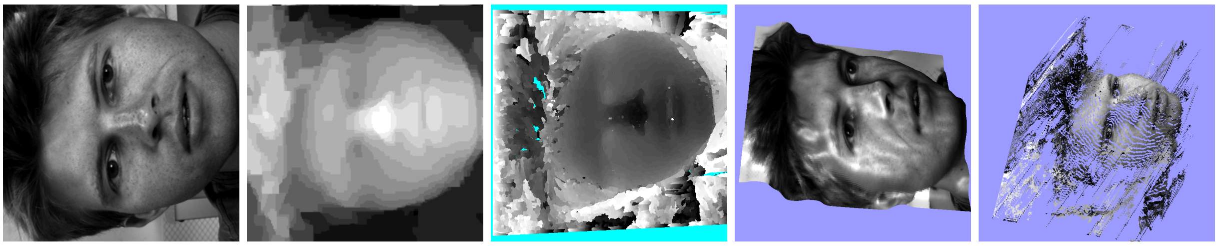


Figure 2: 3D image

The disparity map is coherent. We recognize the hair, the nose, the eyes, and the background. The 3D image corresponds to reality, even though the nose seems a little flat.

3 Comparaison with region growing methods



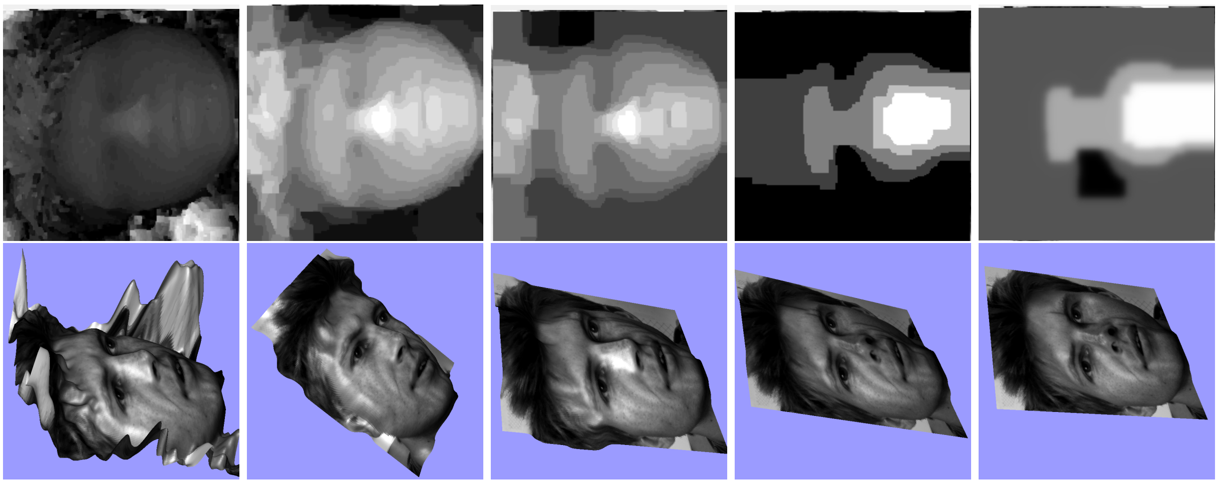
(a) Picture 1 (b) GC Disparity (c) Seeds Disparity (d) GC 3D (e) Seeds 3D

(a) Picture 2 (b) GC Disparity (c) Seeds Disparity (d) GC 3D (e) Seeds 3D

Comparing the Graph-Cut method with the region growing method, we observe that the Graph Cut method gives better result in terms of disparity map and 3D reconstitution for both images. Indeed,we can see with the first picture that the background is better depicted by the Graph-Cut disparity map, which gives better result for the 3D around the face. In particular the image is much smoother for Graph-Cut. In terms of computation time, the Seeds method returns a complete (after propagation) disparity map after 4 minutes 42 seconds for the first image, whereas the Graph-Cut method returns a disparity map in less than 10 seconds for the same image 4 times reduced.

4 Influence of parameters

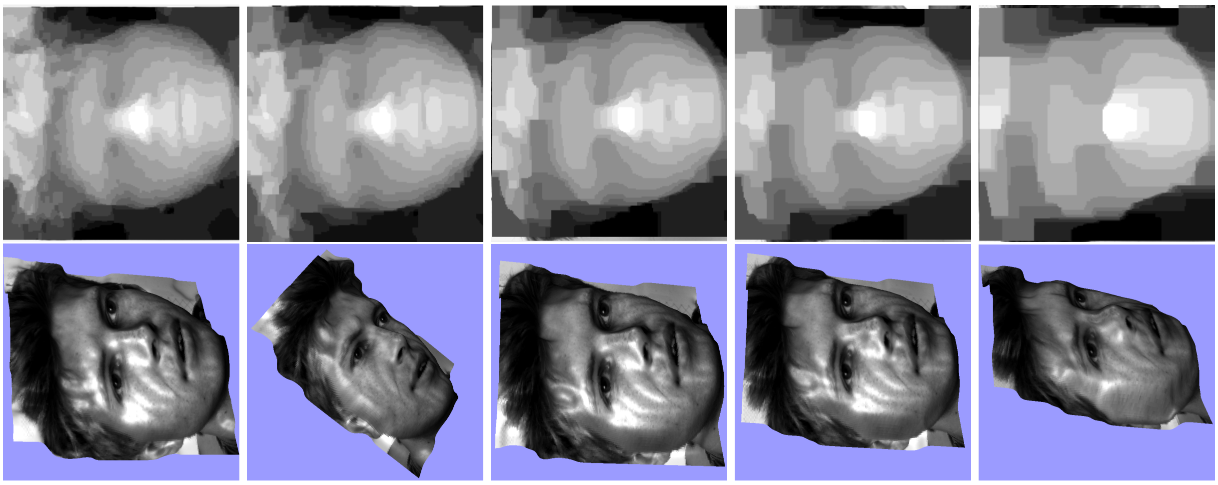
4.1 Parameter λ : smoothing term



(a) $\lambda = 1$ (b) $\lambda = 2$ (c) $\lambda = 5$ (d) $\lambda = 10$ (e) $\lambda = 15$

We can see that as we increase λ, the 3D becomes smoother. This is due to the fact that increasing this parameter increases the cost of attributing different disparities to neighbor pixels. We can clearly see that for a large λ, the disparity map has very few disparity values.

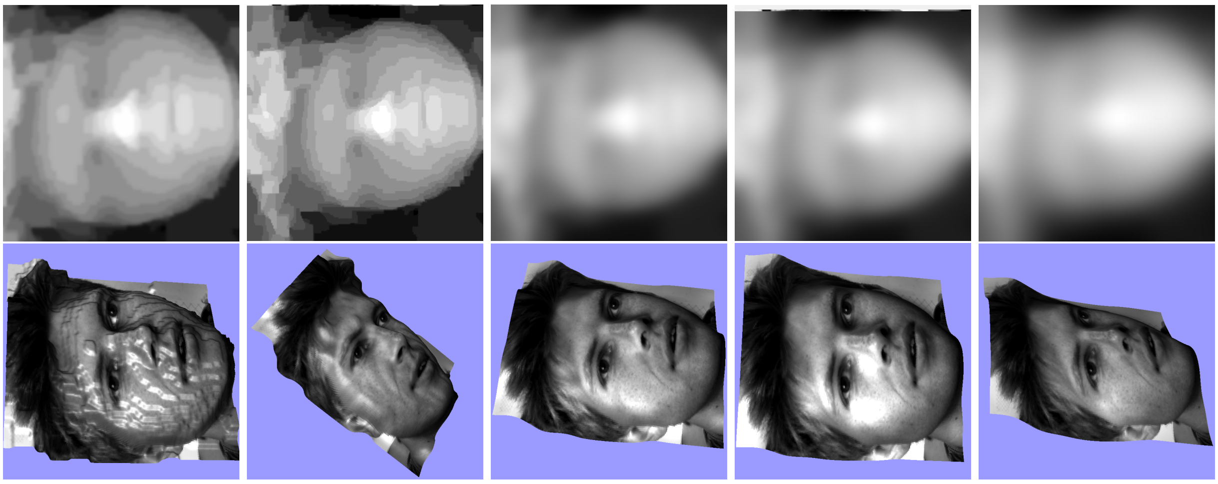
4.2 Parameter n : NCC neighborhood size



(a) $n = 1$ (b) $n = 3$ (c) $n = 7$ (d) $n = 11$ (e) $n = 20$

We can see that as we increase the NCC neighborhood size n, the 3D face becomes more ”cubic”. Indeed, as n increases, we look at averaged cross correlations over bigger patches, and therefore pixels close to the border of the face will be associated to a smaller disparity, whereas the ones at a distance n from the border will keep their big disparity, which creates lines on the face with disparities on each side of the line very different.

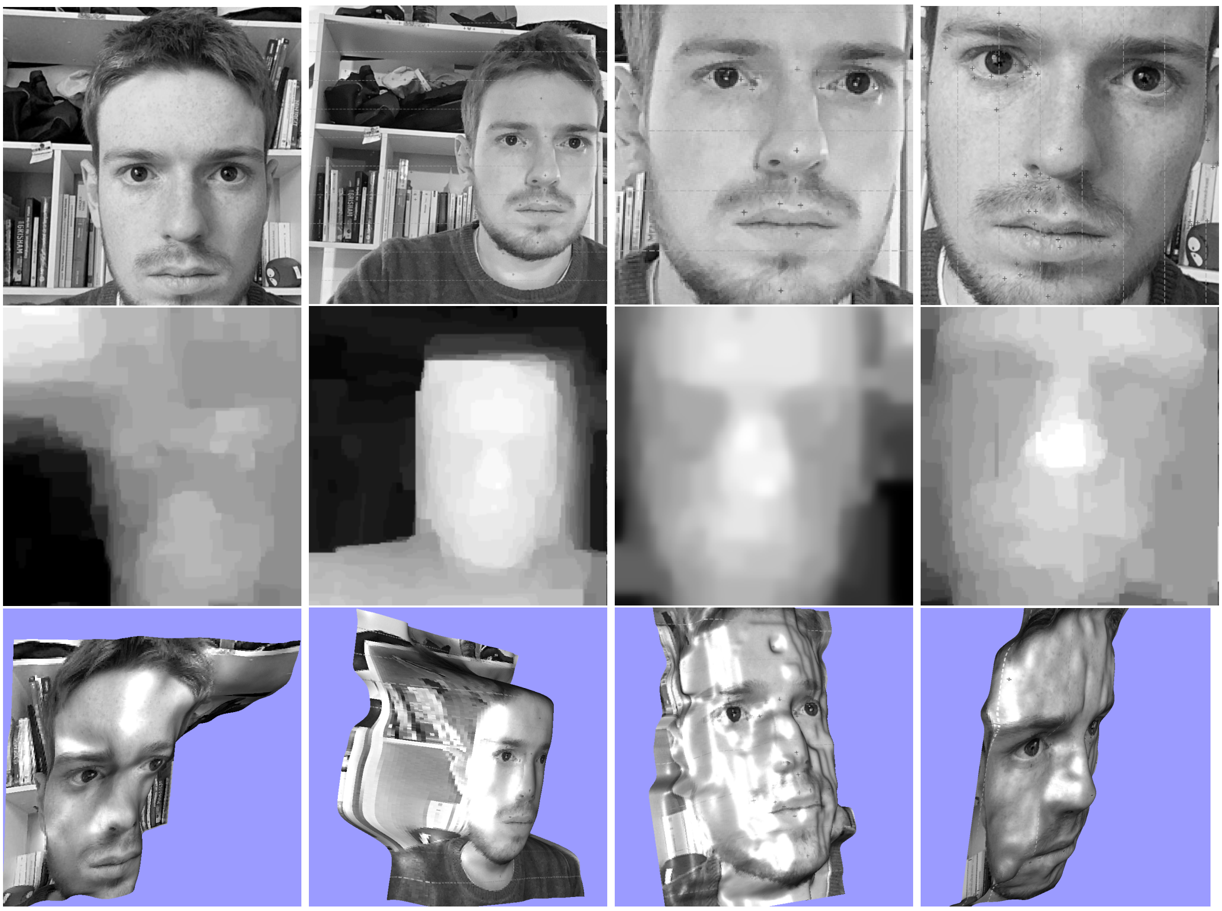
4.3 Parameter σ : Gaussian blur



(a) $\sigma = 1$ (b) $\sigma = 3$ (c) $\sigma = 7$ (d) $\sigma = 9$ (e) $\sigma = 13$

Increasing the blur parameter σ makes the picture smoother.

5 Results with different images



(a) Not rectified (b) Rectified 1 (c) Rectified 1 with zoom (d) Rectified 2

I tried the program with different pairs of images of my face, which led to interesting observations. First, we can see in case a) that images with no epipolar rectification leads to a very bad disparity map and a very bad 3D face. In case b) I rectified the image with IPOL, but took a picture with a very far away background. We can see that with a far background, the method has difficulty recognizing the differentces disparities of my face, which leads to very binary disparity and 3D maps. In case c) I took the same image as in case b) but zoomed closely on my face to give less importance to the background, and the results seem better. In case d) I took a picture only of my face with no background, so the picture has better resolulation than in case c), and it gave the best results. I notice however that the result is not as good as the pictures given by the professor. This can be either because of resolution, either because disparities for these pictures are perfectly adapted (10-55).

6 Conclusion

In this report I compared a Graph Cut method and a region growing method for computing disparity maps and 3D reconstruction. I played with some of the parameters of the Graph Cut program to exhibit their influence on the results. Finally, I enjoyed using the Graph Cut program to produce 3D reconstruction on different images. This eneabled me to understand how to best condition our images in order for the program to process them optimally.