COMBINING TWO METHODS TO ACCURATELY ESTIMATE DENSE DISPARITY MAPS

Agustín Salgado and Javier Sánchez
Computer Science Department, University of Las Palmas de Gran Canaria
35017 Las Palmas de Gran Canaria, Spain

Keywords: disparity map, optical flow, stereoscopic vision.

Abstract: The aim of this work is to put together two methods in order to improve the solutions for the problem of 3D geometry reconstruction from a stereoscopic pair of images. We use a method that we have developed in recent works which is based on an energy minimisation technique. This energy yields a partial differential equation (PDE) and is well suited for accurately estimating the disparity maps. One of the problems of this kind of techniques is that it depends strongly on the initial approximation. For this reason we have used a method based on graph–cuts which has demonstrated to obtain good initial guess.

1 INTRODUCTION

In this paper we have put together two methods for computing disparity maps. The first one is based on graph–cuts energy minimisation (Kolmogorov et al., 2001), (Boykov et al., 2004). This method has demonstrated to give good results in integer precision which is enough for a set of applications. If we are looking for better accuracy then it is necessary to use a different technique. In this case we use a method that we have developed recently and which is described in paper (Alvarez et al., 2002). We have also implemented a similar method for optical flow estimation which is explained in (Alvarez et al., 2000). These methods are based on an energy minimisation approach. When we minimize the energy we obtain a system of PDEs which are then embedded into a gradient descent method to obtain the solution. One of the problems of these methods is that they need a good initial approximation in order to obtain a precise solution. In previous works we have always used a correlation based technique to compute this approximation. Comparing graph–cuts and correlation based methods the first one provides more stable solutions.

In this paper we show that the combination of graph–cuts and PDE based methods improves the accuracy of the solution with respect to other initial approximation techniques such as correlation. In the experimental results we compare numerically the different approaches through a synthetic sequence of a cylinder and also we show several results for a real stereoscopic pair of images – the Tsukuba sequence.

2 GRAPH-CUTS METHOD

The minimum cut/maximum flow algorithms on graphs emerged as an increasingly useful tool for exact or approximate energy minimisation in low-level vision. Stereo is a classical vision problem where graph-based energy minimisation methods have been successfully applied. The goal of stereo is to compute the correspondence between pixels of two or more images of the same scene obtained by cameras with slightly different view points. Any stereo images of multi-depth objects contain occluded pixels. The presence of occlusions adds significant technical difficulties to the problem of stereo.

The energy function for a configuration f is of the form

$$E(f) = E_{data}(f) + E_{occ}(f) + E_{smooth}(f) \quad (1)$$

The three terms here include

- a data term E_{data}, which results from the differences in intensity between corresponding pixels;
- a occlusion term E_{occ}, which imposes a penalty for making a pixel occluded; and

Salgado A. and Sánchez J. (2005).
COMBINING TWO METHODS TO ACCURATELY ESTIMATE DENSE DISPARITY MAPS.
In *Proceedings of the Second International Conference on Informatics in Control, Automation and Robotics - Robotics and Automation*, pages 210-214
DOI: 10.5220/0001178102100214

- a smoothness term E_{smooth}, which makes neighboring pixels in the same image tend to have similar disparities.

3 ENERGY BASED METHOD

The method we use in this paper is energy based. It has the following features:

- We consider a weakly calibrated stereoscopic system. The stereoscopic system is not calibrated and only the knowledge of the so-called fundamental matrix is known.
- This method addresses the problem of accurately determining the dense disparity map while regularizing it along the contours of the gray level image and inhibiting smoothing across the image discontinuities.
- We apply a multi–resolution scheme in order to avoid convergence to irrelevant minima.

The energy function that we propose for 3D geometry reconstruction is as follows:

$$E(\lambda) = \int \left(I_l(\mathbf{x}) \, - \, I_r(\mathbf{x} + \mathbf{h}(\lambda(\mathbf{x}))\right)^2 d\mathbf{x} \\ + C \int \Phi\left(\nabla I_l, \nabla \lambda\right) \, d\mathbf{x}. \quad (2)$$

In this case we have a matching function, $\mathbf{h}$, that depends on a scalar function, λ. This scalar function represents the displacement of pixels on the epipolar lines. In this case $\Phi\left(\nabla I_l, \nabla\lambda\right) = \nabla\lambda^t \cdot D\left(\nabla I_l\right) \cdot \nabla\lambda$, $D\left(I_l \nabla\right)$ is a regularized projection matrix perpendicular to ∇I_l,

$$D\left(\nabla I_l\right) = \frac{1}{|\nabla I_l|^2 + 2\upsilon^2} \cdot \left\{ \begin{bmatrix} \frac{\partial I_l}{\partial y} \\ \frac{-\partial I_l}{\partial x} \end{bmatrix} \begin{bmatrix} \frac{\partial I_l}{\partial y} \\ \frac{-\partial I_l}{\partial x} \end{bmatrix}^t + \upsilon^2 Id \right\} \quad (3)$$

where *Id* denotes the identity matrix. This projection has been introduced by Nagel and Enkelmann in the context of optical flow estimation.

After minimising this energy and applying a gradient descent method we obtain the following diffusion–reaction PDE:

$$\frac{\partial\lambda}{\partial t} = C \operatorname{div}\left(D\left(\nabla I_l\right) \, \nabla\lambda\right) \\ + \left(I_l(\mathbf{x}) - I_r^{\lambda}(\mathbf{x})\right) \cdot \\ \left(\frac{-b\left(\frac{\partial I_r}{\partial x}\right)^{\lambda}(\mathbf{x})}{\sqrt{a^2+b^2}} + \frac{a\left(\frac{\partial I_r}{\partial y}\right)^{\lambda}(\mathbf{x})}{\sqrt{a^2+b^2}} \right) \quad (4)$$

The details of this method could be found in paper (Alvarez et al., 2002).

4 COMBINING GRAPH-CUTS AND STEREOFLOW METHOD

In this section, we explain how the graph-cuts (kz2) and the previous explained PDE (stereoFlow) methods work together for estimating the dense disparity map. The graph-cuts method labels the image obtaining a disparity map in integer precision. The stereoFlow method obtains a disparity map in float precision. To improve the perfomance of our method, we do not apply the graph-cuts method in the input pair of images. Using a pyramidal approach we scale the image *"n"* times. The number of scales is a parameter defined by user.

The basic idea of embedding our method in a pyramidal approach is as follows: we replace the images I_l and I_r by $I_l^\sigma := Z(I_l)$ and $I_r^\sigma := Z(I_r)$, where $Z(\ldots)$ is the zoom operator. Thus, we do a $2X$ zoom over each image. We start with a large initial scale σ_0. Next, we choose a number of scales $\sigma_n < \sigma_{n-1} < \cdots < \sigma_0$ and for each scale σ_i we do a zoom. When we reach last scale (σ_n), we compute the disparity λ_{σ_n} with kz2 or with a correlation based technique. Thus, we have an initial approximation. Next, we compute the disparity λ_{σ_i} as the asymptotic state of the above PDE with initial data $\lambda_{\sigma_{i+1}}$. So, the disparity of I_l and I_r is defined by λ_{σ_0}. In Fig. 1, we see an example how this algorithm works.

Both correlation–based and graph-cuts methods spend much CPU time to compute the disparity maps. As we can see in Fig. 1, the kz2 is applied at the smallest size of the images, so we assure that it is carried out faster than at larger images. Then the stereoFlow technique is applied in the rest of the scales.

5 EXPERIMENTAL RESULTS

In this section we present a comparison between the graph–cuts stereo method (kz2) and the combination of our method (stereoFlow) with different initial approximation (such as kz2 or correlation based technique). We have used two datasets in our tests: a stereoscopic pair from the University of Tsukuba (Fig. 2) and a stereoscopic pair of a synthetic cylinder (Fig. 6).

In paper (Kolmogorov et al., 2001), the head of Tsukuba was used to show the results obtained with graph-cuts stereo method (kz2) in comparison with similar methods. We have used the same dataset to

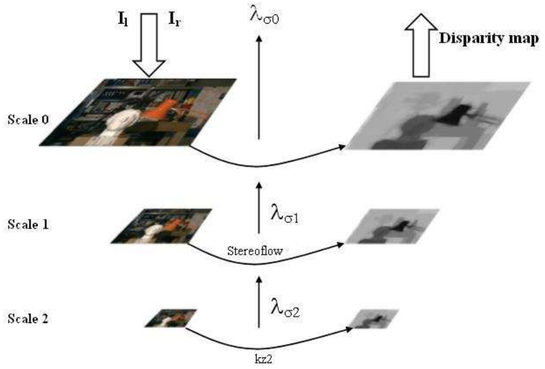

Figure 1: Combination of kz2 and stereoFlow algorithms

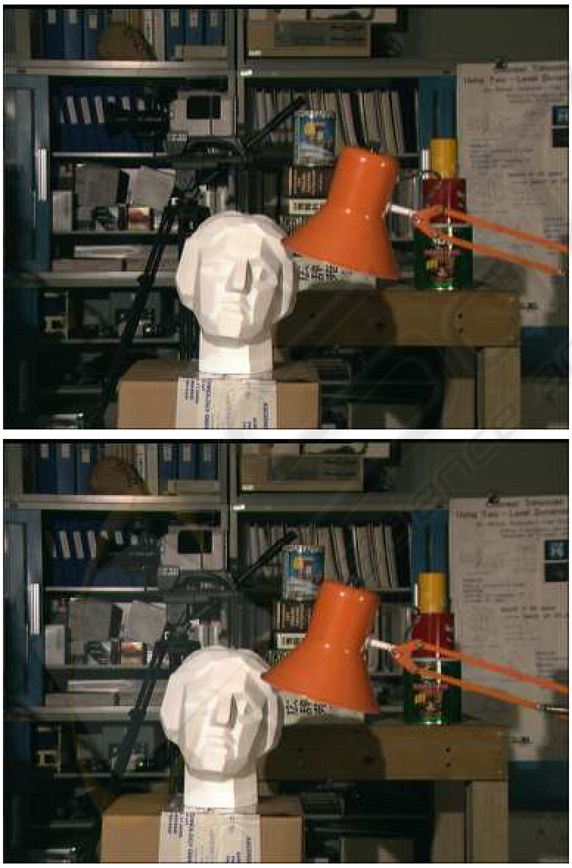

Figure 2: Stereoscopic pair for the Tsubuka sequence.

Figure 3: Disparity map obtained through kz2.

Figure 4: Disparity map obtained through correlation + stereoflow.

show how our algorithm improves the initial approximation given by kz2 (with/without scales).

The number of zooms defined by user depends on the scene motion. An overzooming (or overscaled) gives us a bad initial approximation, so our method converges to irrelevant local minima. In Fig. 3 we can see the disparity map obtained by kz2. Both correlation based techniques and graph-cuts methods spend much CPU time so we must decide between a few or large number of scales.

We have tested our method using the synthetic cylinder dataset to compare its accuracy with different initial approximations. For the initial value we consider two possibilities. The first one is to use the result of a simple classic method for estimating the disparity, for instance a correlation based technique. The second one is to use a graph-cuts method which gives us a disparity map in integer precision.

We have computed the euclidean error between the

Figure 5: Disparity map obtained through kz2 + stereoflow.

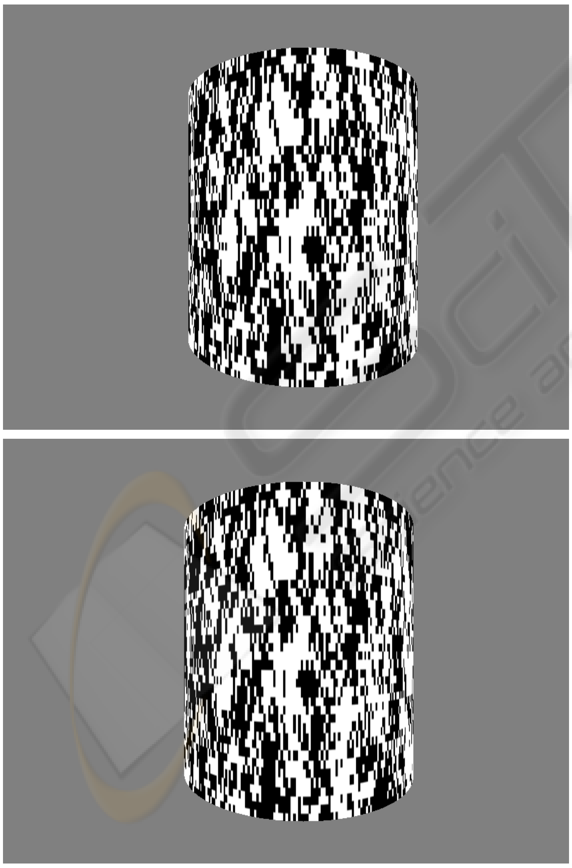

Figure 6: Stereoscopic pair for the synthetic cylinder.

Table 1: Euclidean error obtained with cylinder images for various tests.

Euclidean error		
Method	**Disparity range**	
	Scale=0	**Scale**=1
kz2	0.24	0.24
Corr+StereoFlow	0.22	0.21
kz2+StereoFlow	0.20	0.19

output of our algorithm and the ideal disparity map, to see the accuracy of our method. In the table 1 we show the results obtained by the combination of a correlation–based method with stereoFlow, the result for the kz2 and the result for the combination of kz2 and stereoFlow.

From these results we may appreciate that the combination of correlation and stereoFlow method gives better results than the kz2 method and that the combination of kz2 and stereoFlow improves the solution of the correlation–based one.

In this table we compare the solution for two different configurations: In the first one the pyramidal scheme is reduced to only one scale (Scale = 0) and in the second one we use two scales (Scale = 1). If we compare them, we may conclude that the use of the pyramidal approach improves the solutions for both methods (correlation + stereoFlow and kz2 + stereoFlow) and that the result for kz2 + stereoFlow is still better than using the correlation–based method.

In figures 8 and 10, we see the visual results for the synthetic cylinder obtained for each method. Looking at the figures the result obtained with the latter is smoother and more accurate. All the experimental results are improved with the combined method and in most cases the improvement is greater than a 16% for the euclidean error for the same scale.

6 CONCLUSIONS

In this work we have combined two different techniques on disparity maps estimation in order to obtain more accurate and reliable solutions. We have used a pixel precision method based on graph–cuts as initialization for another method based on PDEs. The latter depends on an initial approximation which is supported by the former one. The solution we obtain is in float precision and the accuracy is considerably improved. We have compared the combination of the PDE and graph–cuts with the combination of the PDE and a correlation–based method. We may conclude that the use of the kz2 at the first stage provides better results than the correlation method.

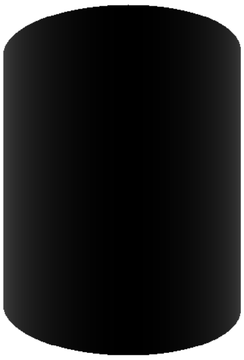

Figure 7: Ideal disparity map for the cylinder images.

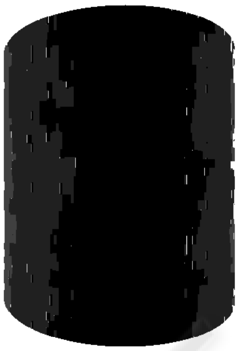

Figure 8: Disparity map obtained through kz2.

Figure 9: Disparity map obtained through correlation + stereoFlow.

Figure 10: Disparity map obtained through kz2 + stereoFlow.

REFERENCES

L. Alvarez, R. Deriche, J. Sánchez and J. Weickert. Dense Disparity Map Estimation Respecting Image Discontinuities: A PDE and Scale-Space Based Approach. Journal of Visual Communication and Image Representation, 13, pp. 3-21. 2002.

L. Alvarez J. Weickert and J. Sánchez. Reliable Estimation of Dense Optical Flow Fields with Large Displacements. International Journal of Computer Vision, 39, pp. 41-56, 2000.

V. Kolmogorov and R. Zabih, Computing Visual Correspondence with Occlusions using Graph Cuts. In International Conference on Computer Vision (ICCV), July 2001.

Y. Boykov and V. Kolmogorov, An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision. In IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), September 2004.