Towards Real-Time Stereoscopic Image Rectification for 3D Visualization

Tony Marrero Barroso   Aubrey K. Dunne   John Mallon   Paul F. Whelan

CIPA (*Centre for Image Processing and Analysis*)
Dublin City University, Ireland
www.cipa.dcu.ie

Introduction

- Stereo Vision systems are used in order to obtain *Three Dimensional* instantaneous depth information from the world, a human's binocular vision is such an example.

- In order to merge both left and right views and extract the depth from a scene, a human's eyes must be aligned. The same principle applies when 3D stereo images are generated using a pair of images taken from different viewpoints of the scene.
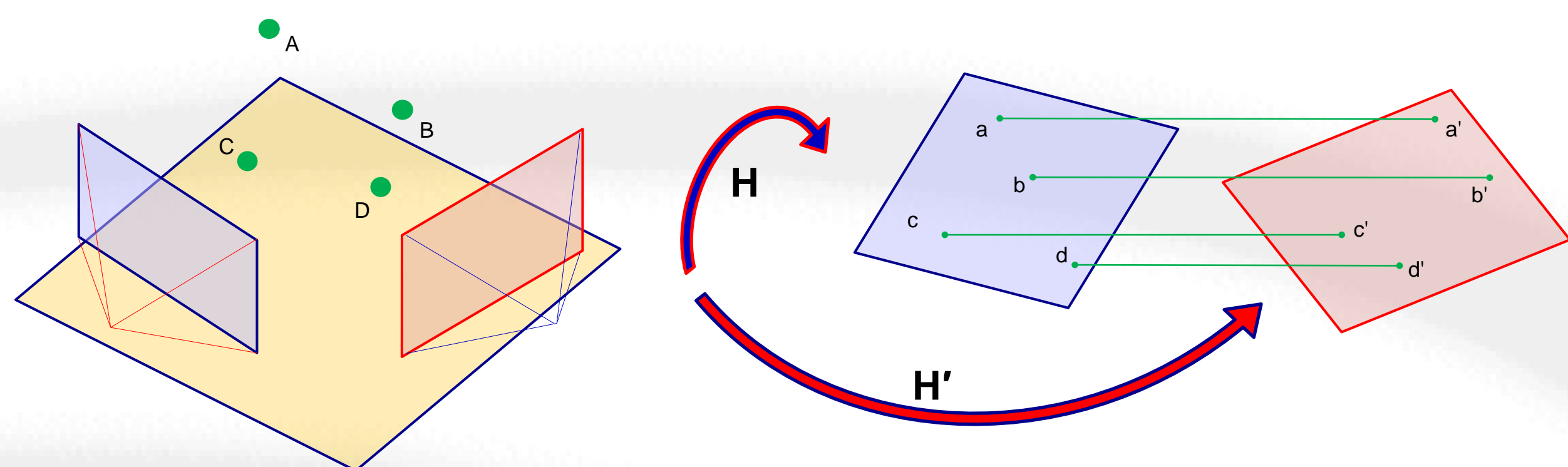

Figure 1: The Stereo Rectification Process. Left hand diagram shows the initial triangulation scene. The second image illustrates the image rectification results, all corresponding features become horizontally aligned.

- In order to ensure that corresponding objects are aligned in both left and right images, a technique called *Stereo Rectification* is utilised. Planar homographies (*H and H'*) are applied to warp both images accordingly (*see Fig. 1*).

- The *Epipolar Geometry* between both images is analysed using matched corresponding feature points, H and H' can then be derived from the *Fundamental Matrix* (*F*).

- In order to realistically recreate the depth illusion inherent in stereoscopic 3D visualisation, the warping *distortion must be minimised* during the rectification process.

- This work presents a distortion minimisation strategy that is *significantly more computationally efficient* than the previous state of the art method found in the literature.

- Of specific interest is the synthesis of stereoscopic 3D video content from a monocular video sequence, which can also be relevant to *Augmented Reality* applications.
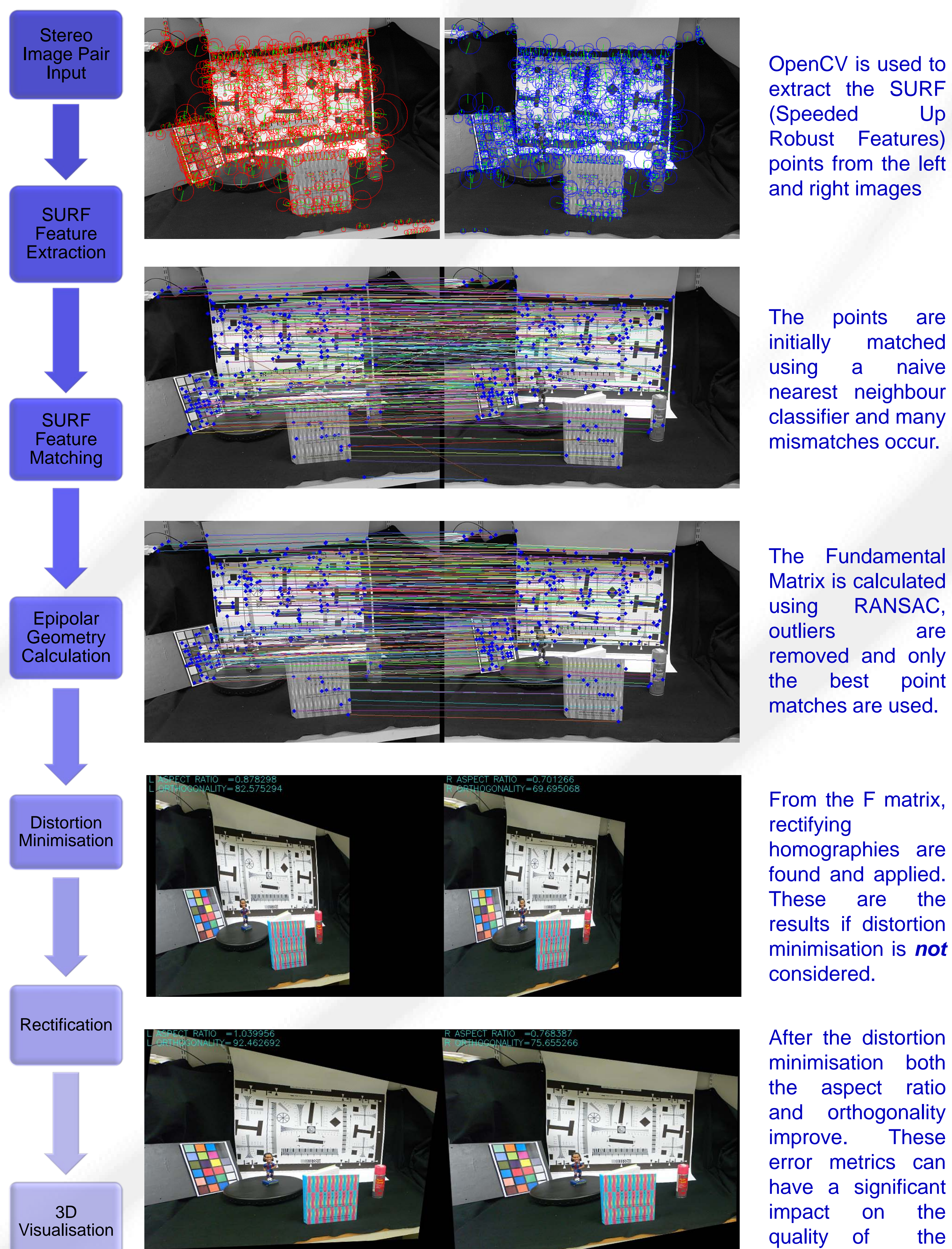
Pipeline

Figure 2: Image sequence of our implemented rectification pipeline.

OpenCV is used to extract the SURF (Speeded Up Robust Features) points from the left and right images

The points are initially matched using a naive nearest neighbour classifier and many mismatches occur.

The Fundamental Matrix is calculated using RANSAC, outliers are removed and only the best point matches are used.

From the F matrix, rectifying homographies are found and applied. These are the results if distortion minimisation is **not** considered.

After the distortion minimisation both the aspect ratio and orthogonality improve. These error metrics can have a significant impact on the quality of the visualisation.

Distortion Minimisation

- The implemented stereo rectification approach *minimises horizontal distortions* on the rectified image pairs by analysing the *Jacobian* (*J*) of the pixel's transformation across all local neighbourhoods of the image.

- The method attempts to make the singular values (σ_1, σ_2) of these Jacobians as close to one as possible, and thus reduce distortions by minimising local scale re-sampling effects. This in turn ensures that the aspect ratio and orthogonality of the rectified images remain near the ideal values of 1 and 90° respectively.

- To find the appropriate variables of the rectifying homographies, the following cost function is minimised. The state of the art technique minimises this cost function using the unconstrained *Nelder-Mead* simplex search method.

$$f(\hat{a}_{11}, \hat{a}_{12}) = \sum_{i=1}^{N}\left[(\sigma_1 - 1)^2 + (\sigma_2 - 1)^2\right]$$

- Our approach facilitates the use of more efficient non-linear optimisation techniques by the development of analytical constraints that relate the cost function to the unknown variables. These techniques can then avail of full analytical derivatives to optimise the search.

Results

- Three standard minimisation algorithms (*Gradient Descent*, *Gauss-Newton* and *Levenberg-Marquardt*) were implemented in C to test the performance and robustness of the developed mathematical constraints. A rigorous comparison was made between our implementations and the results obtained using the Nelder-Mead method of Numerical Recipes in C.

- To put the results in perspective, the popular L-M minimisation method of Minpack was also compared. The tests compared the efficiency of the minimisation along with the accuracy.

Table 1: Comparison of the average results of all the minimisation tests performed.

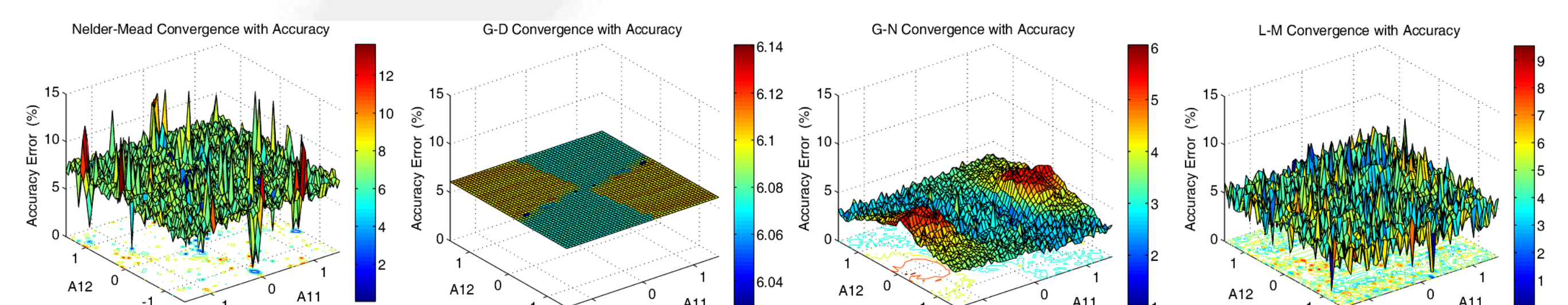| Algorithm		Accuracy Error (%)	Cost Function Evaluations	Time (ms)
Gradient Descent	(G-D)	6.125	7	0.065
Gauss-Newton	(G-N)	3.625	17	0.113
Levenberg-Marquardt	(L-M)	4.920	16	0.110
Minpack's L-M		4.180	54	0.141
Nelder-Mead		7.133	60	0.785

Figure 3: *Accuracy of Convergence*; plots comparing the accuracy of the Nelder-Mead method with our three implemented minimisation algorithms. With the developed mathematical constraints, these three algorithms were able to converge *more consistently and accurately* than the Nelder-Mead search method under varying conditions.
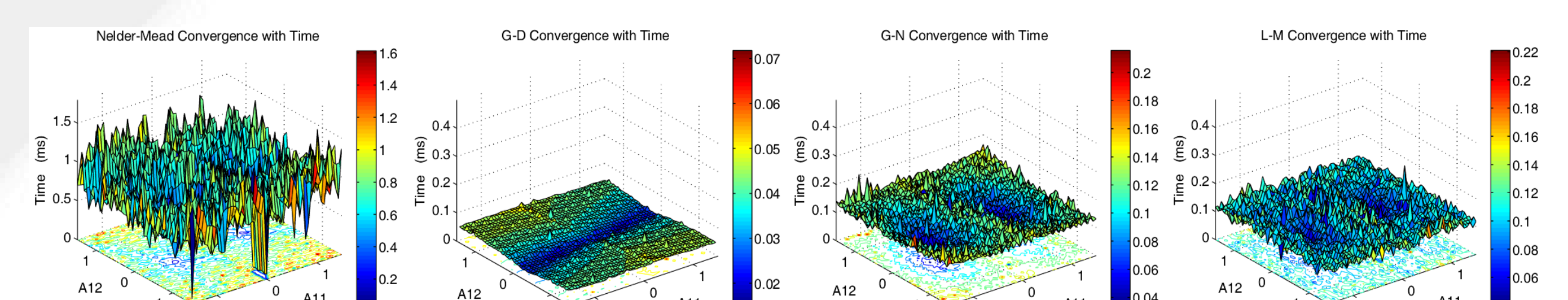

Figure 4: *Time of Convergence*; plots comparing the computational efficiency of the Nelder-Mead method with our three implemented minimisation algorithms. The Nelder-Mead search method had a *significantly inferior performance* under most of the tested conditions. The G-D method in some conditions was even up to 1200% times faster.
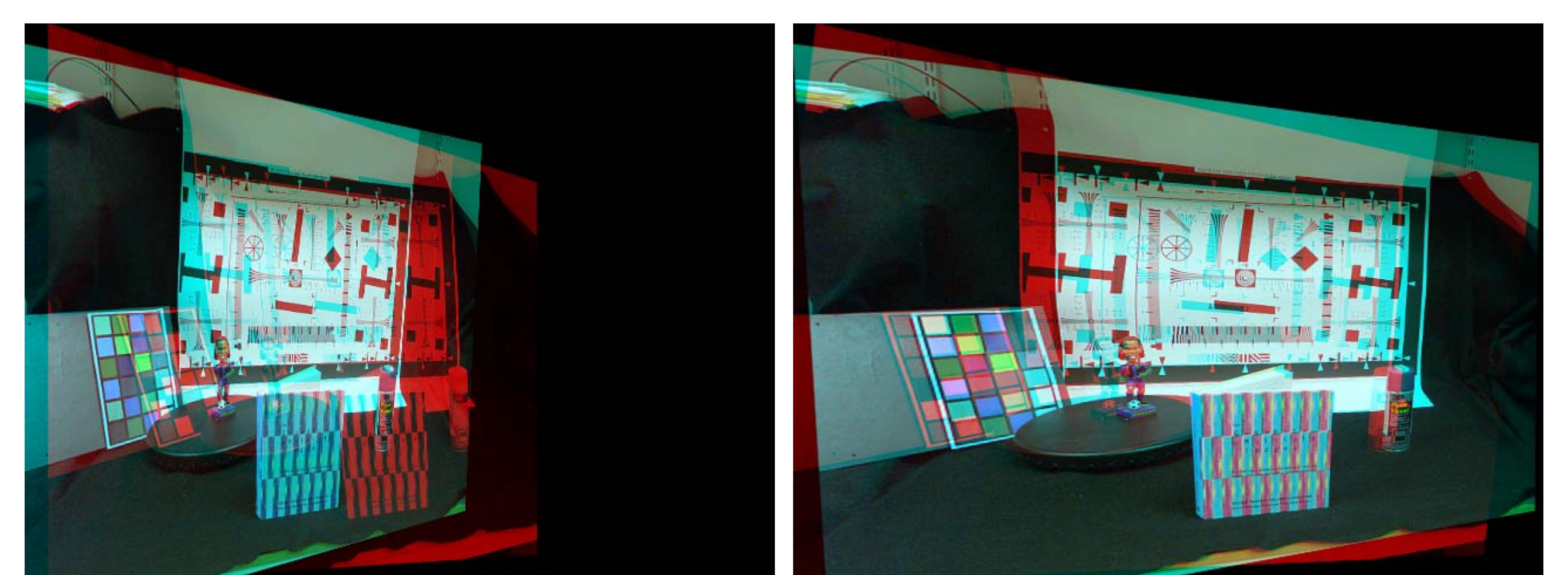

Figure 5: Stereoscopic 3D Anaglyph image **without** distortion minimisation. The left and right images cannot be merged effectively due to the horizontal shearing distortions.
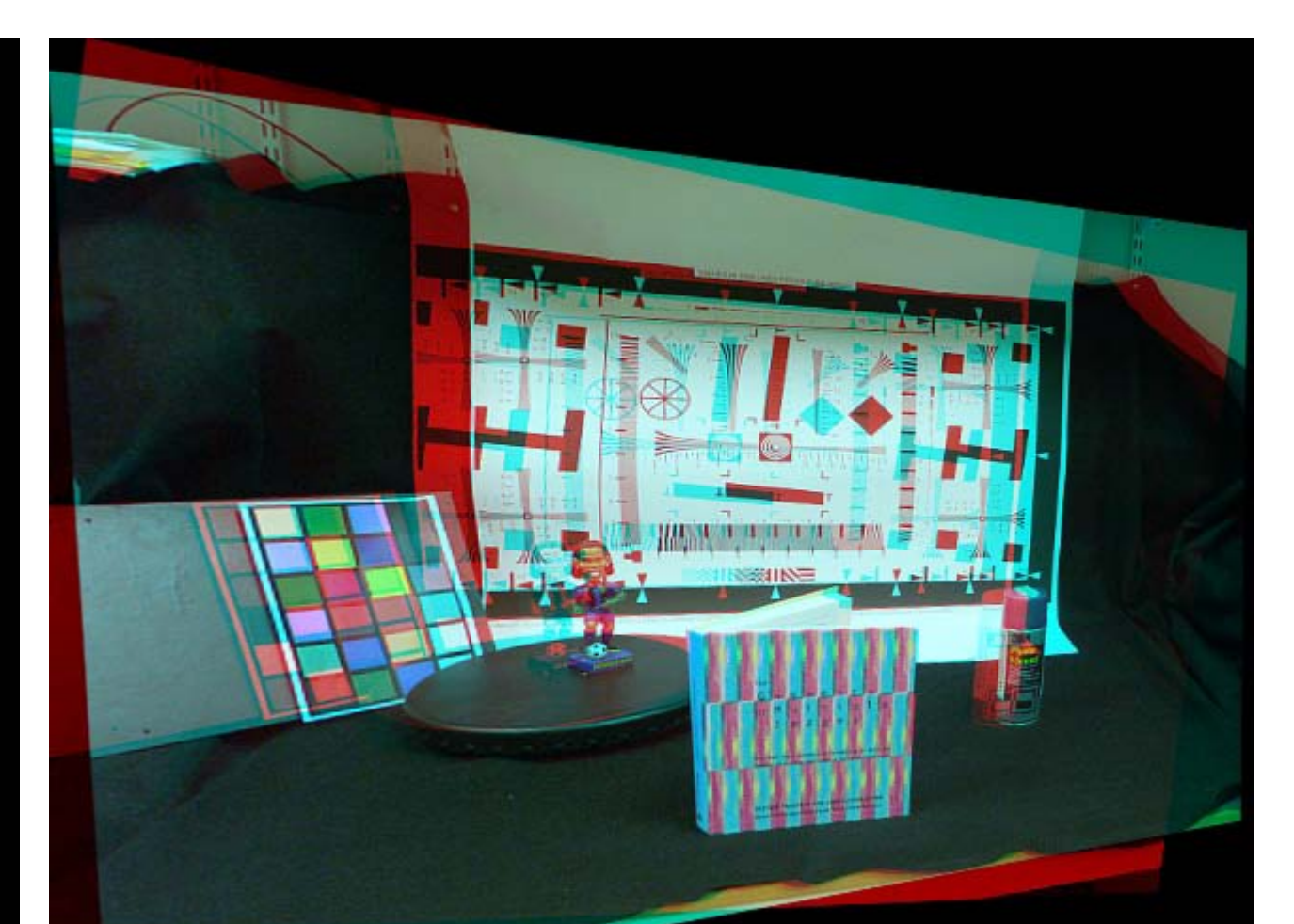
Figure 6: Stereoscopic 3D Anaglyph image **with** distortion minimisation. All objects are also aligned, but the improvement in the 3D visualisation that occurs after the distortion minimisation is considerable.

Acknowledgements